



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA ELEKTROTECHNIKY  
A KOMUNIKAČNÍCH TECHNOLOGIÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

**ÚSTAV RADIOELEKTRONIKY**

DEPARTMENT OF RADIO ELECTRONICS

**SROVNÁNÍ STANDARDŮ VP9 A H.265 PRO  
KÓDOVÁNÍ VIDEOSEKVENČÍ**

COMPARISON OF VP9 AND H.265 ENCODING STANDARDS

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

Vítězslav Fendrich

**VEDOUCÍ PRÁCE**

SUPERVISOR

Ing. Ladislav Polák, Ph.D.

**BRNO 2017**

# Bakalářská práce

bakalářský studijní obor **Elektronika a sdělovací technika**

Ústav radioelektroniky

**Student:** Vítězslav Fendrich

**ID:** 164201

**Ročník:** 3

**Akademický rok:** 2016/17

## NÁZEV TÉMATU:

### Srovnání standardů VP9 a H.265 pro kódování videosekvencí

#### POKyny PRO VYPRACOVÁNÍ:

V teoretické části práce detailně prostudujte vlastnosti kodeků VP9 a H.265 pro kódování videosekvencí v moderních multimediálních službách. S využitím dostupných referenčních implementací jednotlivých kodeků vypracujte postupy pro komprimaci videosekvencí s vysokým rozlišením a vysokou kvalitou. Uvažujte formáty videa typické pro mobilní aplikace (relativně nízké rozlišení, nižší snímková rychlost), stejně tak jako formáty UHDTV. Prostudujte a vyberte vhodné objektivní a subjektivní metriky pro hodnocení kvality komprimovaných videosekvencí.

V experimentální části práce komprimujte vhodné videosekvence pomocí uvažovaných kodeků. Srovnajte výkonnost jednotlivých uvažovaných implementací (závislost kvality na kompresním poměru) pomocí vybraných metrik. Zaměřte se na detailní analýzu a provnání objektivních a subjektivních výsledků. Navrhněte laboratorní úlohu pro srovnání kodeků VP9 a H.265 pro kódování krátkých videosekvencí. Připravte vzorové vypracování navržené úlohy.

#### DOPORUČENÁ LITERATURA:

[1] KUFA, J., POLAK, L., KRATOCHVIL, T. HEVC/H.265 vs. VP9 for Full HD and UHD video: Is there any difference in QoE? In Proceedings of 58th International Symposium ELMAR 2016. Zadar (Croatia), 2016, pp. 51-55.

[2] ITU-R Recommendation BT.500-13. Methodology for the subjective assessment of the quality of television pictures. Geneva (Switzerland): ITU, Jan. 2012.

**Termín zadání:** 6.2.2017

**Termín odevzdání:** 30.5.2017

**Vedoucí práce:** Ing. Ladislav Polák, Ph.D.

**Konzultant:**

**prof. Ing. Tomáš Kratochvíl, Ph.D.**  
*předseda oborové rady*

#### UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

# ABSTRAKT

Tato bakalářská práce se zabývá srovnáním vlastností standardů VP9 a HEVC (H.265), vyvinutých pro kódování videa zejména ve vysokém rozlišení. Teoretická část práce se zaměřuje na podrobný popis jednotlivých standardů a také na popis vybraných objektivních a subjektivních metrik pro hodnocení kvality videa. Praktická část práce se zabývá vyhodnocením kvality různých videosekvencí, komprimovaných pomocí kodeků VP9 a H.265, vybranými objektivními metrikami a subjektivní metodou hodnocení. Dále je také popsán postup komprese vybraných videosekvencí a popis realizace subjektivních testů v laboratorních podmínkách. Získané výsledky jsou přehledně vyhodnoceny a diskutovány. Současně byla navržena laboratorní úloha k ověření uvedených poznatků.

# KLÍČOVÁ SLOVA

H.265, HEVC, VP9, ACR, PSNR, SSIM, VIFP, VQM, komprese videa

# ABSTRACT

This bachelor thesis deals with a comparison of the features of video coding standards VP9 and HEVC (H.265), developed to encoding videos mainly in high resolution. Theoretical part of this work is focused on the presentation and description of both VP9 and H.265 video codecs. Furthermore, objective and subjective metrics for video quality assessment are briefly described. Practical part of this work deals with an evaluation of the quality of video sequences, encoded by VP9 and H.265, with objective metrics and subjective metric. The description of the encoding process using VP9 and H.265 is also included. Proceeding of subjective tests in laboratory conditions is also described. Achieved results are clearly presented and discussed.

# KEYWORDS

H.265, HEVC, VP9, ACR, PSNR, SSIM, VIFP, VQM, video compression

FENDRICH, V. *Srovnání standardů VP9 a H.265 pro kódování videosekvencí* . Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2017. 51 s, 14 s. příloh. Vedoucí bakalářské práce Ing. Ladislav Polák, Ph.D..

# PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Srovnání standardů VP9 a H.265 pro kódování videosekvencí“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne 29. května 2017

.....

(podpis autora)

# PODĚKOVÁNÍ

Děkuji vedoucímu této bakalářské práce Ing. Ladislavu Polákovi, Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování této bakalářské práce.

V Brně dne 29. května 2017

.....

(podpis autora)

Experimentální část této diplomové práce byla realizována na výzkumné infrastruktuře  
vybudované v rámci projektu CZ.1.05/2.1.00/03.0072

**Centrum senzorických, informačních a komunikačních systémů (SIX)**  
operačního programu Výzkum a vývoj pro inovace.

# OBSAH

<b>Seznam obrázků</b>	<b>viii</b>
<b>Seznam tabulek</b>	<b>x</b>
<b>Úvod</b>	<b>1</b>
<b>1 Kodek VP9</b>	<b>2</b>
1.1 Popis kodéru .....	2
1.2 Dělení snímků .....	2
1.3 Intrapredikce .....	3
1.4 Interpredikce .....	4
1.5 Transformační kódování .....	5
1.6 Entropické kódování .....	5
1.7 Rekonstrukční filtr .....	5
<b>2 HEVC (H.265)</b>	<b>6</b>
2.1 Popis kodéru .....	6
2.2 Stromová struktura.....	7
2.3 Intrapredikce .....	8
2.4 Interpredikce .....	9
2.5 Entropické kódování .....	9
2.6 Rekonstrukční filtr .....	10
<b>3 Metriky hodnocení kvality videa</b>	<b>11</b>
3.1 Objektivní metriky hodnocení kvality videa .....	11
3.1.1 PSNR .....	11
3.1.2 SSIM .....	12
3.1.3 VQM .....	13
3.1.4 VIF .....	15
3.2 Subjektivní metriky hodnocení kvality videa .....	15
3.2.1 ACR .....	16
3.2.2 DSCQS.....	17
<b>4 Postupy pro komprimaci videosekvencí</b>	<b>19</b>
4.1 Referenční implementace VP9 .....	19
4.1.1 Kompilace programu .....	19

4.1.2	Příkazy .....	19
4.2	Referenční implementace HEVC (H.265) .....	20
4.2.1	Ovládání kodeku .....	20
<b>5</b>	<b>Srovnání kompresní účinnosti jednotlivých standardů</b>	<b>22</b>
5.1	Metoda srovnání .....	22
5.2	Použitý software .....	22
5.3	Videosekvence .....	23
5.3.1	Přehled videosekvencí .....	24
5.3.2	Informace o prostorové a časové aktivitě (SI a TI) .....	25
5.4	Subjektivní testy .....	27
5.5	Výsledky .....	29
5.5.1	Sekvence Tall Buildings .....	29
5.5.2	Sekvence Campfire .....	34
5.5.3	Sekvence Fountains .....	39
5.5.4	Sekvence Woods .....	44
5.6	Korelační analýza výsledků .....	50
<b>Závěr</b>		<b>51</b>
<b>Literatura</b>		<b>52</b>
<b>Seznam symbolů, veličin a zkratk</b>		<b>54</b>
<b>A Laboratorní úloha</b>		<b>55</b>
A.1	Návrh laboratorní úlohy .....	55
A.2	Vzorové vypracování laboratorní úlohy .....	62
<b>B DVD</b>		<b>69</b>



# SEZNAM OBRÁZKŮ

Obrázek 1-1 Příklad dělení snímku do jednotlivých bloků (na základě [3]) .....	3
Obrázek 1-2 Jednotlivé režimy intrapredikce (na základě [1]).....	4
Obrázek 2-1 Blokové schéma kodéru HEVC (na základě [4]).....	7
Obrázek 2-2 Rozdělení PU bloků (na základě [3]).....	8
Obrázek 2-3 Jednotlivé režimy intrapredikce (převzato z [3]) .....	9
Obrázek 3-1 Schéma postupu testování metodou ACR.....	16
Obrázek 3-2 Hodnotící formulář DSCQS (na základě [1]) .....	17
Obrázek 5-1 Diagram postupu srovnání sekvencí .....	22
Obrázek 5-2 Informace o prostorové a časové aktivitě .....	26
Obrázek 5-3 Ishihariho pseudoizochromatické tabulky (převzato z [26]) .....	27
Obrázek 5-4 Testovací pracoviště.....	29
Obrázek 5-5 Tall Buildings UHD průměrné hodnoty PSNR v závislosti na bitové rychlosti .....	30
Obrázek 5-6 Tall Buildings UHD hodnoty SSIM v závislosti na bitové rychlosti.....	30
Obrázek 5-7 Tall Buildings UHD hodnoty VQM v závislosti na bitové rychlosti.....	31
Obrázek 5-8 Tall Buildings UHD hodnoty VIFP v závislosti na bitové rychlosti .....	31
Obrázek 5-9 Tall Buildings UHD průměrné hodnoty MOS.....	32
Obrázek 5-10 Tall Buildings SD průměrné hodnoty PSNR v závislosti na bitové rychlosti .....	32
Obrázek 5-11 Tall Buildings SD hodnoty SSIM v závislosti na bitové rychlosti .....	33
Obrázek 5-12 Tall Buildings SD hodnoty VQM v závislosti na bitové rychlosti .....	33
Obrázek 5-13 Tall Buildings SD hodnoty VIFP v závislosti na bitové rychlosti.....	34
Obrázek 5-14 Tall Buildings SD průměrné hodnoty MOS .....	34
Obrázek 5-15 Campfire UHD průměrné hodnoty PSNR v závislosti na bitové rychlosti .....	35
Obrázek 5-16 Campfire UHD hodnoty SSIM v závislosti na bitové rychlosti .....	35
Obrázek 5-17 Campfire UHD hodnoty VQM v závislosti na bitové rychlosti .....	36
Obrázek 5-18 Campfire UHD hodnoty VIFP v závislosti na bitové rychlosti .....	36
Obrázek 5-19 Campfire UHD průměrné hodnoty MOS.....	37
Obrázek 5-20 Campfire SD průměrné hodnoty PSNR v závislosti na bitové rychlosti .	37

Obrázek 5-21 Campfire SD hodnoty SSIM v závislosti na bitové rychlosti .....	38
Obrázek 5-22 Campfire SD hodnoty VQM v závislosti na bitové rychlosti .....	38
Obrázek 5-23 Campfire SD hodnoty VIFP v závislosti na bitové rychlosti.....	39
Obrázek 5-24 Campfire SD hodnoty MOS.....	39
Obrázek 5-25 Fountains UHD Průměrné hodnoty PSNR v závislosti na bitové rychlosti .....	40
Obrázek 5-26 Fountains UHD hodnoty SSIM v závislosti na bitové rychlosti.....	40
Obrázek 5-27 Průměrné hodnoty VQM v závislosti na bitové rychlosti.....	41
Obrázek 5-28 Fountains UHD hodnoty VIFP v závislosti na bitové rychlosti .....	41
Obrázek 5-29 Fountains UHD průměrné hodnoty MOS .....	42
Obrázek 5-30 Fountains SD průměrné hodnoty PSNR v závislosti na bitové rychlosti	42
Obrázek 5-31 Fountains SD hodnoty SSIM v závislosti na bitové rychlosti .....	43
Obrázek 5-32 Fountains SD hodnoty VQM v závislosti na bitové rychlosti .....	43
Obrázek 5-33 Fountains SD hodnoty VIFP v závislosti na bitové rychlosti .....	44
Obrázek 5-34 Fountains SD průměrné hodnoty MOS.....	44
Obrázek 5-35 Woods UHD průměrné hodnoty PSNR v závislosti na bitové rychlosti .	45
Obrázek 5-36 Woods UHD hodnoty SSIM v závislosti na bitové rychlosti .....	45
Obrázek 5-37 Woods UHD hodnoty VQM v závislosti na bitové rychlosti .....	46
Obrázek 5-38 Woods UHD hodnoty VIFP v závislosti na bitové rychlosti.....	46
Obrázek 5-39 Woods UHD průměrné hodnoty MOS .....	47
Obrázek 5-40 Woods SD průměrné hodnoty PSNR v závislosti na bitové rychlosti.....	47
Obrázek 5-41 Woods SD hodnoty SSIM v závislosti na bitové rychlosti.....	48
Obrázek 5-42 Woods SD hodnoty VQM v závislosti na bitové rychlosti.....	48
Obrázek 5-43 Woods SD Průměrné hodnoty VIFP v závislosti na bitové rychlosti.....	49
Obrázek 5-44 Woods SD průměrné hodnoty MOS .....	49

# SEZNAM TABULEK

Tabulka 2-1 Přehled dostupných verzí kódování H.265 (na základě [4]) .....	7
Tabulka 3-1 Kategorie ACR (podle [9]).....	16
Tabulka 4-1 Příkazy syntaxe VP9 (na základě [16]) .....	20
Tabulka 4-2 Příkazy syntaxe HEVC (na základě [19]) .....	21
Tabulka 5-1 Přehled jednotlivých videosekvencí.....	25
Tabulka 5-2 Korelační koeficienty objektivních metrik.....	50

# ÚVOD

Poptávka po poskytování multimediálních služeb, zejména videa, v různých kvalitách se rapidně zvyšuje. Současné bezdrátové systémy mají své limity (např. šířka pásma), a tak není možné přenášet video ve vysokém rozlišení Full High Definition (HD), případně ultra vysokém rozlišení Ultra High Definition (UHD) bez patřičné komprese, na kterou jsou kladeny stále vyšší požadavky. Moderní kódovací a dekódovací algoritmy tak musí efektivně redukovat bitový tok videa při zachování kvality obrazu videa a současně příliš nezatěžovat výpočetní výkon zařízení.

Tato práce je zaměřena na porovnání vlastností dvou v současnosti nejmodernějších standardů s označením VP9 a High Efficiency Video Coding (HEVC/H.265) pro kódování videosekvencí v ultra vysokém rozlišení. Cílem práce je provést srovnání výkonnosti obou uvažovaných standardů, prostřednictvím vybraných metod objektivního a subjektivního hodnocení kvality komprimovaných sekvencí.

Bakalářská práce je členěna celkem do pěti kapitol. První dvě kapitoly jsou věnovány teoretickému rozboru fungování obou kódovacích standardů VP9 a HEVC. Třetí kapitola se zabývá výběrem a představením objektivních a subjektivních metod hodnocení obrazové kvality. V této části jsou popsány plně či částečně referenční objektivní metody. Představeny jsou také dvě různé subjektivní metody, které se odlišují použitím srovnávací referenční sekvence. Ve čtvrté kapitole jsou uvedeny postupy pro komprimaci sekvencí s využitím dostupných referenčních implementací jednotlivých standardů. Poslední, pátá kapitola je věnována samotnému provedení srovnání účinnosti obou standardů, pomocí komprimovaných sekvencí ve vysokém (UHD) a nízkém rozlišení. Kvalita jednotlivých sekvencí byla následně vyhodnocena pomocí uvažovaných objektivních metrik. Současně byly uspořádány subjektivní testy hodnocení kvality v laboratorních podmínkách. Následuje statistické vyhodnocení dat a v závěru porovnání jednotlivých získaných výsledků objektivních a subjektivních testů. Také byla navržena laboratorní úloha, sloužící k ověření uvedených poznatků a taktéž bylo připraveno vzorové vypracování této úlohy.

# 1 KODEK VP9

Kodek s označením VP9 představuje novou generaci videokodeku pro kódování sekvencí ve vysokém rozlišení. Jde o zdokonaleného nástupce předchozí verze VP8, který je součástí širšího projektu s názvem WebM, jenž funguje pod záštitou společnosti Google Inc. Jedná se o otevřený (tzv. open-source) formát, kde základní myšlenkou projektu WebM je vybudovat širokou komunitu vývojářů, kteří se budou podílet na vývoji a zdokonalování kodeku. Finální verze kodeku byla vydána v červnu roku 2013. Díky svému masivnímu rozšíření byl tento kodek zahrnut do multimediální knihovny s názvem libvpx, která implementuje příslušný audio a video kontejner [1].

Potřeba vzniku nové verze kodeku vychází ze stále se zvyšujících nároků uživatelů Internetu na kvalitu přenášeného videa při zachování, nebo ještě lépe snížení, přenosové rychlosti. Prakticky totožné důvody vedly ke vzniku nové verze konkurenčního kodeku s označením H.265, o kterém pojednává následující kapitola této práce.

Jak vyplývá z výše uvedeného, hlavním místem uplatnění VP9 jsou webové aplikace – zejména internetové videoservery (např. YouTube), ale také programy a aplikace pro videokonference. Je také podporován všemi hlavními internetovými prohlížeči. Vzhledem k hlavnímu účelu využití je pro tento kodek nejdůležitější široká podpora napříč různými platformami – osobní počítače, mobilní zařízení, chytré televizory či multimediální centra [2].

## 1.1 Popis kodéru

Kodek VP9 je zaměřen především na video ve vysokém rozlišení, samozřejmostí je tedy podpora standardního rozlišení 720p (1280 x 720 pixelů), dále vysokého rozlišení Full HD, někdy též označovaného jako 1080p (1920 x 1080 pixelů). Po poslední aktualizaci kodek podporuje také stále populárnější velmi vysoké rozlišení 4K, neboli Ultra HD (UHD), s typickým rozlišením 3840 x 2160 pixelů [1].

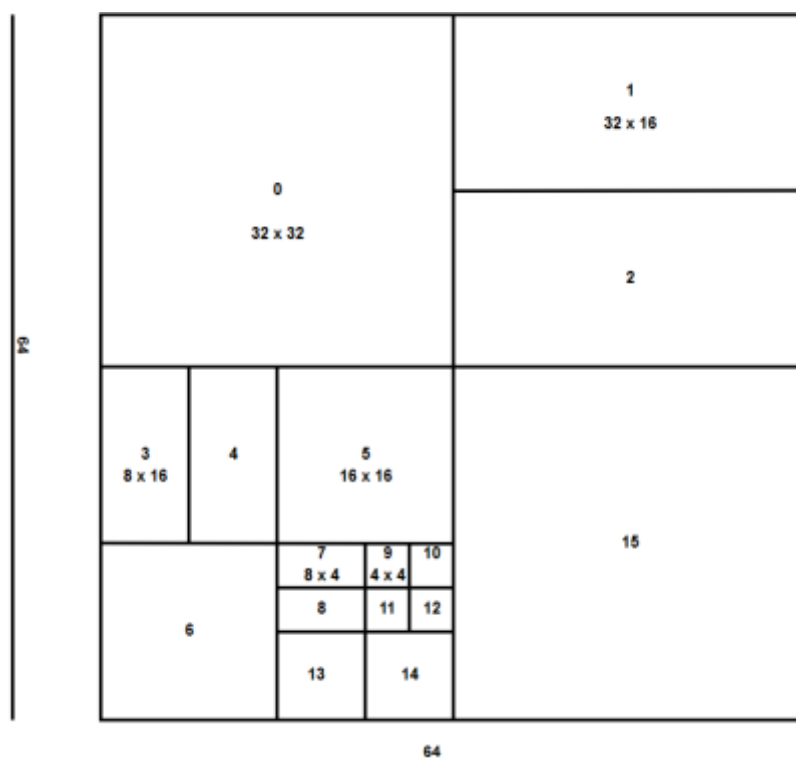
Původní verze kodeku VP9 podporovala pouze jediný typ barevného vzorkování a to YCbCr 4:2:0 s 8-bitovou hloubkou. S postupem času a průběžným vývojem nyní VP9 podporuje čtyři výchozí profily, které rozšiřují možnosti vzorkování i nastavení bitové hloubky. Profil 0 je základním profilem s výše uvedenými parametry, profil 1 přidává podporu vzorkování YCbCr 4:2:2 a 4:4:4. Profily s označením 2 a 3 dále podporují 10-bitovou a 12-bitovou hloubku a dále podporu alfa kanálu (průhlednosti) [2].

## 1.2 Dělení snímků

Ve VP9 je jednou z největších změn oproti kodeku VP8 podstatné zvýšení velikosti blokových oblastí jednotlivých snímků z důvodu zvýšení efektivity kódování. VP9 přinesl takzvané superbloky (SB) o velikosti až 64 x 64 pixelů. Tyto superbloky mohou být dále děleny na menší superbloky o velikosti 32 x 32 pixelů a makrobloky o velikosti 16 x 16 pixelů. Makrobloky mohou být z důvodu zvýšení jemnosti predikce a pro účely

transformace děleny až na nejmenší možnou velikost 4 x 4 pixelů [1]. Dělení obrazu do jednotlivých blokových oblastí názorně demonstruje Obrázek 1-1.

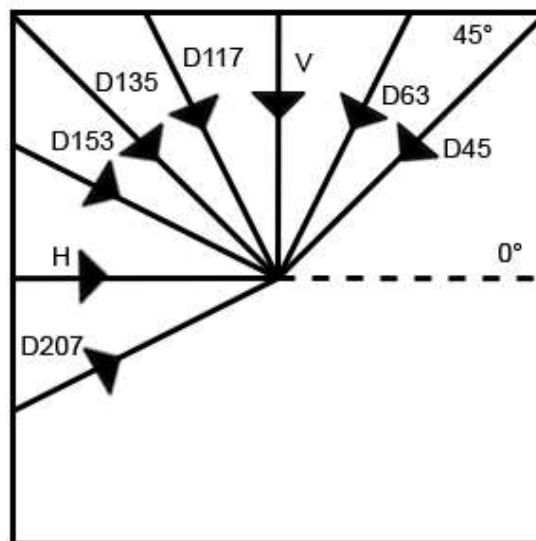
Při zpracovávání videa ve vysoké kvalitě je žádoucí, aby procesy kódování a dekódování videa probíhaly paralelně, a to z důvodu efektivního využití dostupného výpočetního výkonu zařízení. Toto je nutné především u zařízení, která disponují nižším výpočetním výkonem. Z toho důvodu podporuje VP9 takzvaný *frame\_parallel\_mode*, kdy se jednotlivé snímky rozdělí na dlaždice a v jednotlivých hlavičkách rámců bitového proudu jsou uvedeny značky, sloužící k paralelnímu dekódování datového proudu [2].



Obrázek 1-1 Příklad dělení snímku do jednotlivých bloků (na základě [3])

### 1.3 Intrapredikce

Kodek VP9 podporuje sadu 10 módů intrapredikce pro různé velikosti bloků, které se pohybují v rozmezí 4 x 4 až 32 x 32 pixelů (Obrázek 1-1). Definovány jsou predikce horizontální (H\_PRED), vertikální (V\_PRED), predikce s označením True Motion (TM\_PRED) a dále sada 6 diagonálních predikcí o různých úhlech, přičemž označení daného predikčního módu označuje současně úhel predikce. Například predikční mód D45 je 45 stupňovou interpolací hodnot vzorků. Jednotlivé úhly i směr interpolace jednotlivých intrapredikčních módů názorně ilustruje Obrázek 1-2.



Obrázek 1-2 Jednotlivé režimy intrapredikce (na základě [1]).

## 1.4 Interpredikce

VP9 podporuje čtyři interpredikční módy pro velikost bloků od minimální velikosti 4 x 4 až po nejvyšší možnou velikost 64 x 64 pixelů. Pro pohybovou kompenzaci používá kodek primárně jeden pohybový vektor, ale nabízí také možnost složené predikce pomocí dvou pohybových vektorů na jednom bloku, jehož výsledkem jsou dva predikované vzorky, které jsou následně zprůměrovány [1].

Pohybový vektor ukazuje na jeden ze tří možných referenčních snímků, které se označují jako *Last frame*, neboli jeden z předchozích kódovaných snímků, dále *Golden frame*, neboli referenční snímek uložený v bufferu, který se využívá například při stříhu scény na jinou a poté zase zpět, a dále *Alternate reference frame*, který se na rozdíl od ostatních nemusí vůbec zobrazit a může sloužit pouze jako referenční za účelem zvýšení přesnosti predikce [1].

U kodeku VP9 jsou pohybové vektory predikovány z uspořádaného seznamu možných referenčních pohybových vektorů. Tento seznam „kandidátů“ na pohybové vektory je vytvořen až z osmi obklopujících bloků, které sdílí stejný referenční obraz. Pokud tento proces nenaplní seznam referenčních vektorů, pak jsou obklopující bloky prohledány znovu, nicméně tentokrát již bez ohledu na referenční obraz.

Po přiřazení danému bloku je následně kódován jeden ze čtyř možných módů pohybového vektoru. Mód s označením `NEW_MV` použije první zápis ve výše uvedeném seznamu spolu s pohybovým vektorem delta, který je vyslán v datovém toku. Oproti tomu mód s označením `NEAREST_MV` použije pouze první zápis v seznamu bez dalších úprav. Obdobné je to i u módu s označením `NEAR_MV`, který použije druhý zápis v seznamu bez úprav. Poslední mód `ZERO_MV` používá pohybový vektor o hodnotě (0,0) [3].

## 1.5 Transformační kódování

VP9 používá tři základní druhy transformací: diskrétní kosinovou transformaci (DCT), asymetrickou diskrétní sinovou transformaci (ADST) a Walsh-Hadamardovu transformaci (WHT) [3]. Použití konkrétního typu transformace závisí na použitém režimu predikce. Transformace DCT se nejčastěji používá pro bloky o velikosti  $4 \times 4$  až  $32 \times 32$  pixelů, a to při intrapredikci. Pro interpredikci je nejčastěji využívána transformace ADST často v kombinaci s DCT, kdy vzniká takzvaná hybridní transformace pro bloky o velikosti  $4 \times 4$  až  $16 \times 16$  pixelů. Transformace typu WHT se používá pouze pro bezeztrátové kódování videa, a to v kombinaci s bloky o velikosti  $4 \times 4$  pixelů. Velikost bloku určeného pro transformaci je zde zpravidla volena jako nejvyšší možná, tedy odpovídá velikosti predikčního bloku.

## 1.6 Entropické kódování

Kodek využívá výhradně aritmetického entropického kódování, při kterém je vybraný symbol z  $n$  prvkové abecedy reprezentován prostřednictvím sestaveného binárního stromu, který se tak skládá z celkového počtu  $n - 1$  hlavních uzlů. Binární aritmetický kodér pak pracuje na každém z hlavních uzlů s 8-bitovou přesností [1].

S ohledem na dynamické změny obrazu videí je kódování plně adaptivní. Většina starších kódovacích standardů (např. MPEG2) používá systém zpětné adaptace, kdy jsou pravděpodobnostní hodnoty aktualizovány dle počtu symbolů během kódování či dekódování. VP9 však disponuje možností obousměrné adaptace, která přináší možnost adaptace dle následujících klíčových snímků (key frame). Kombinace obou těchto metod přispívá k nárůstu efektivity kódovacího a dekódovacího procesu během zpracování videa [1].

## 1.7 Rekonstrukční filtr

Jak již bylo uvedeno výše, VP9 podporuje relativně velké množství kombinací velikostí jednotlivých bloků. Z tohoto důvodu je hlavním cílem rekonstrukčního filtru vyhlazení případných ostrých hran bloků v obraze. Takzvaný „deblock filter“ nejprve zahlučuje vertikální hrany bloků v celém snímku a až poté hrany horizontální. Z důvodu možného velkého rozdílu velikostí bloků disponuje VP9 filtry o šířce 16, 8, 4 a 2 pixely, přičemž jejich použití je plně adaptivní [3].

VP9 v rámci rekonstrukčního filtru také disponuje detekováním plochosti snímku (flatness detector), který detekuje ploché či statické části snímku, a na základě toho upravuje velikost použitého filtru spolu s jeho intenzitou použití [1].



## 2 HEVC (H.265)

Kodek s názvem H.265, označovaný jako High Efficiency Video Coding (HEVC), je prozatím nejnovějším standardem pro kompresi videosekvencí, který je orientován zejména na videa ve vysoké kvalitě. Kodek byl vyvinut organizacemi International Telecommunication Union (ITU-T), Video Coding Experts Group (VCEG) a ISO/IEC Moving Picture Experts Group (MPEG), které za účelem spolupráce na vývoji tohoto kódovacího standardu vytvořily společnost s názvem Joint Colaborative Team on video Coding (JCT-VC). První verze tohoto kodeku byla vydána v lednu roku 2013 [4].

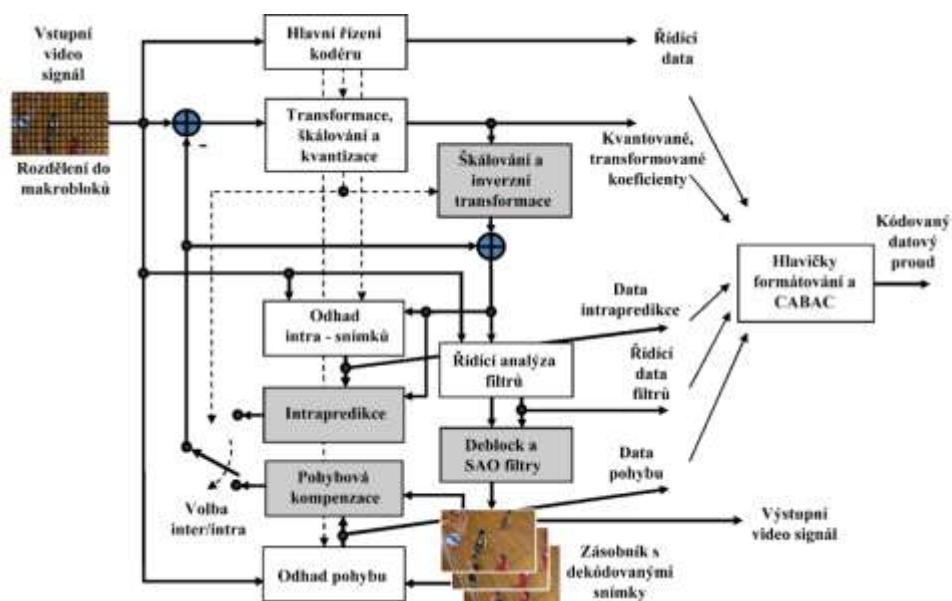
Principiálně kodek vychází ze svého velmi úspěšného předchůdce s označením H.264 Advanced Video Coding (AVC) od stejných tvůrců [5]. HEVC je však primárně zaměřen na zpracování videa ve vysokém rozlišení, a to jak Full HD, tak UHD. Zajímavostí je nově implementovaná podpora videa v rozlišení 8K UHD (8192 x 4320 pixelů). Hlavním cílem při vývoji nového kodeku tedy bylo zvýšení podporovaného maximálního rozlišení a zvýšení kompresní účinnosti tak, aby spolu se skokovým zvýšením rozlišení nedocházelo ke shodnému zvyšování datového proudu. Dalším cílem pak bylo maximálně efektivní využití výpočetních prostředků, zejména prostřednictvím paralelně probíhajících procesů. Již podle prvních provedených testů [6] dosahuje HEVC téměř 50% úspory bitového toku oproti svému předchůdci, a to při zachování shodné obrazové kvality.

### 2.1 Popis kodéru

HEVC je hybridní blokově orientovaný kódovací standard, který je založen na prostorové transformaci. Vstupní snímky jsou nejprve rozděleny do stejně velkých bloků označených jako makrobloky. Proces komprese a následného dekódování pak pracuje s každým jednotlivým makroblokem zvlášť. Tyto makrobloky jsou dále rozděleny do menších bloků, mezi kterými následně probíhá snímková predikce. Blokové schéma, které názorně demonstruje princip činnosti HEVC je uvedeno na Obrázek 2-1.

První verze kodeku HEVC podporovala nativně tři základní profily s označením *Main*, *Main10* a *Still picture*. Nejčastěji je využíván základní profil *Main*, který podporuje vzorkování YCbCr 4:2:0 s 8-bitovou barevnou hloubkou. Profil *Main10* měl shodné parametry, ale 10-bitovou hloubku. Profil s označením *Still picture* byl speciálním, spíše doplňkovým profilem, sloužícím pro zobrazení statického obrazu [4].

Druhá vydaná verze kodeku přidala podporu 21 nových kódovacích profilů, které podstatně zvýšily variabilitu možností vzorkování, a také barevné hloubky o velikosti až 16 bitů. Základní přehled jednotlivých profilů nabízí Tabulka 2-1.



Obrázek 2-1 Blokové schéma kodéru HEVC (na základě [4])

Tabulka 2-1 Přehled dostupných verzí kódování H.265 (na základě [4])

Verze 1			Verze 2						
Název	Main	Main10	Main12	Main 4:2:2 10	Main 4:2:2 12	Main 4:4:4	Main 4:4:4 10	Main 4:4:4 10	Main 4:4:4 16
Barevná hloubka [bit]	8	8 - 10	8 - 12	8 - 10	8 - 12	8	8 - 10	8 - 12	8 - 16
Vzorkování	4:2:0	4:2:0	4:2:0	4:2:0/ 4:2:2	4:2:0/ 4:2:2	4:2:0/ 4:2:2/ 4:4:4	4:2:0/ 4:2:2/ 4:4:4	4:2:0/ 4:2:2/ 4:4:4	4:2:0/ 4:2:2/ 4:4:4

## 2.2 Stromová struktura

Koncept makrobloků u HEVC je představován prostřednictvím vytváření kódových stromových jednotek CTU (Coding Tree Unit), které jsou složeny z kódových stromových bloků CTB (Coding Tree Blocks), a to pro jasový i barevný signál. CTU může být o velikosti 16 x 16, 32 x 32 nebo 64 x 64 pixelů. Pro představu, starší standard H.264 podporoval CTU o velikosti pouze 16 x 16 pixelů [5]. Zvýšení velikosti CTU vedlo ke zvýšení efektivity dělení jednotlivých bloků při zpracovávání videa o vysokém rozlišení. Takto vzniklá úspora datového proudu činí přibližně 16 % [4].

Blok CTB se dále dělí na kódové bloky CB (Coding Blocks), které jsou určeny strukturou kvadrantového stromu. Minimální velikost bloku pro jasový signál je 8 x 8 pixelů. V případě formátu vzorkování 4 : 2 : 0 je pak doplněn dvěma barevnými CB, každý o velikosti 4 x 4 pixely. Společně pak tyto bloky tvoří takzvanou kódovou jednotku CU (Coding Unit). Tyto CU bloky jsou nejnižšími jednotkami ve struktuře kvadrantového stromu. CU bloky však dále obsahují predikční bloky PU (Prediction

Unit) a transformační bloky TU (Transformation Unit) [7].

Prvek TU je syntaxe odpovědná za ukládání transformačních dat. Povolené velikosti TU jsou následující:  $32 \times 32$ ,  $16 \times 16$ ,  $8 \times 8$  a  $4 \times 4$  pixelů, s ohledem na velikost matice DFT. Oproti tomu PU je prvek syntaxe, který ukládá predikční informace jako například zvolený úhel intrapredikce nebo použitý pohybový vektor interpredikce. Jednotka CU pak může obsahovat až čtyři jednotky PU. Stavy, podle kterých se blok CU dělí na jednotlivé PU bloky, mohou být následující:  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$ ,  $N \times N$ ,  $2N \times nU$ ,  $2N \times nD$ ,  $nL \times 2N$ , kde  $2N$  reprezentuje rozdělenou velikost bloku CU. Tyto stavy názorně ilustruje obrázek 2.2. V režimu intrapredikce je povoleno pouze rozdělení  $2N \times 2N$  [3].

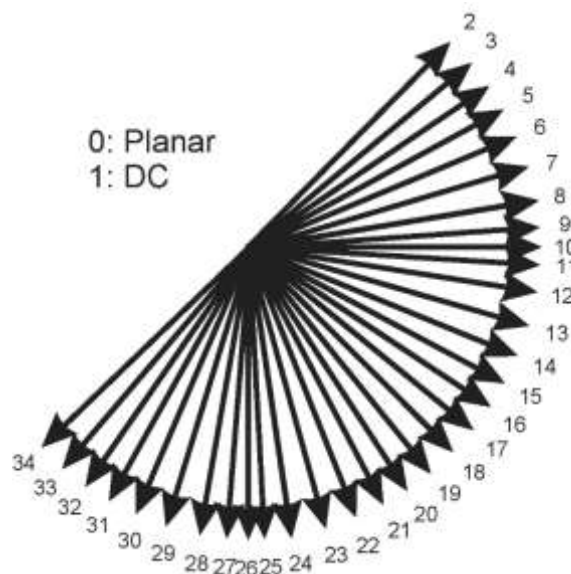


Obrázek 2-2 Rozdělení PU bloků (na základě [3]).

## 2.3 Intrapredikce

Kodek HEVC podporuje tři základní režimy vnitrosnímkové (intra) predikce: planární, Direct Component (DC) a dále 33 směrových režimů predikce. Pro porovnání, starší verze H.264 disponovala pouze 8 směrovými režimy predikce [4]. Celkem tedy HEVC disponuje 35 režimy intrapredikce a umožňuje použití TU bloků o velikostech  $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$  a  $32 \times 32$  pixelů.

Vůbec nejjednodušším typem intrapredikce je režim DC, kde se k predikci používá střední hodnota vzorků z řady nacházející se nad blokem a také po stranách bloku. Naopak nejnáročnější mód z hlediska využití výpočetních prostředků je režim planární, který je tvořen dvoudimenzionální lineární interpolací. Pro úhlový režim intrapredikce jsou prováděny lineární interpolace dostupných bodů ve stanoveném směru. Hodnota velikosti úhlu predikce závisí na úhlu, který svírá vertikální osa se směrem interpolace [2]. Pro interpolace jdoucí po směru hodinových ručiček nabývá kladných hodnot, proti směru naopak záporných. Jednotlivé směry intrapredikce názorně ilustruje Obrázek 2-3.



Obrázek 2-3 Jednotlivé režimy intrapredikce (převzato z [3])

## 2.4 Interpredikce

Obdobně jako H.264 i HEVC pracuje se dvěma referenčními seznamy označenými L0 a L1. Každý z těchto seznamů může obsahovat až 16 referenčních vzorků, nicméně maximální počet unikátních obrazů je 8. To znamená, že některé obrazové vzorky budou v seznamu opakovaně. Toto následně usnadňuje predikci ze stejného obrazu s rozdílnými váhami predikce [2].

Predikce uvnitř snímků probíhá u HEVC prostřednictvím pohybového vektoru. Během predikce je vytvářen seznam vhodných pohybových vektorů a následně je vybrán nejvhodnější kandidát pomocí přiděleného indexu, který je zakódován v datovém proudu. K dispozici jsou dva režimy pro predikci pomocí pohybového vektoru, a to „merge“ a „Advanced Motion Vector Prediction“ (AMVP). Pro každý blok PU rozhoduje kódér o volbě vhodného módu a tuto volbu následně označí značkou, která je taktéž zakódována do datového proudu. Predikce AMVP používá kódování pohybového vektoru delta a může vytvořit jakoukoliv hodnotu pohybového vektoru [2].

## 2.5 Entropické kódování

Kodek HEVC používá adaptivní binární aritmetické kódování, tzv. CABAC (Context Adaptive Binary Arithmetic Coding), na úrovni bloků CTU [4]. Při srovnání s H.264 došlo ke značnému zjednodušení inicializačního procesu a ke zredukování počtu proměnných přibližně o polovinu. Syntaxe datového proudu je navržena tak, aby části kódu, které je možné ignorovat, byly sdruženy co nejvíce u sebe a nedocházelo tak ke zbytečnému zpomalování algoritmu. Kódování CABAC je však pouze sekvenční proces.

Z toho důvodu není možné využívat jakékoliv paralelní procesy, které by jeho zpracování urychlily. To je také jeden z důvodů proč je pro kompresi videa pomocí HEVC potřeba více času ve srovnání s H.264. Stejně tak je s ohledem na množství komplexních procesů, které probíhají při kódování a dekódování videa, požadován vyšší výpočetní výkon [4]. Dle vzájemných srovnání dosahuje výpočetní náročnost HEVC až trojnásobku oproti H.264 [6].

## 2.6 Rekonstrukční filtr

HEVC disponuje dvěma základními filtry pro zpětné zpracování obrazu: Deblock Filter (DBF) a Sample – Adaptive Offset (SAO) [4]. Použití obou těchto filtrů je volitelné a je možné je deaktivovat.

DBF, neboli filtr pro potlačení blokové struktury, se používá pro zjemnění a zamaskování viditelných hran v obrazu coby pozůstatek rozdělení obrazu do makrobloků a jejich následného kódování. Filtr nejprve potlačuje všechny vertikální hrany v obraze a následně zpracuje ty horizontální, přičemž tato filtrace se provádí na všech blocích až do minimální velikosti 8 x 8 pixelů. Pokud je tedy video kódováno s makrobloky o velikosti 4 x 4 pixely, pak filtr DBF není aktivní. Algoritmus filtru je plně adaptivní a sílu filtrování přizpůsobuje kolísání pixelových hodnot poblíž ohraničení jednotlivých bloků.

Jako sekundární filtr je následně použit SAO [4], který provádí nelineární amplitudové mapování prostřednictvím vyhledávací tabulky na úrovni CTB bloků. Filtr proběhne jednou na každém pixelu CTB bloku. Pro každý CTB blok má filtr na výběr ze čtyř hodnot offsetu, které nabývají hodnot od -7 do 7, pro 8-bitovou videosekvenci. Kodér tyto hodnoty volí tak, aby se zdrojový i výchozí obraz co nejvíce shodovaly.

## 3 METRIKY HODNOCENÍ KVALITY VIDEA

Metody, které hodnotí kvalitu videa, se dají rozdělit na dvě základní skupiny, a to na objektivní a subjektivní. Účelem objektivních metrik je matematické vyhodnocení kvality posuzovaných sekvencí prostřednictvím daných algoritmů, které se liší podle dané metriky. Jejich cílem je přiblížit se výsledkům ze subjektivního hodnocení kvality videa. Výsledky z objektivních metrik jsou však ve srovnání se subjektivním hodnocením získány v mnohem kratším čase. Z toho důvodu jsou uvedené metriky často implementovány do řídicích systémů zpracovávajících video. Při subjektivním posuzování kvality videa dochází k přímému hodnocení kvality videosekvence určitým počtem pozorovatelů. Následně je použitím daných statistických metod vyhodnocen výsledný index kvality (například Mean Opinion Score – MOS). Výhodou subjektivních metod je vysoká míra spolehlivosti výsledků, zatímco nevýhodou je jejich časová náročnost. Dále je nutné zajistit dostatečný a reprezentativní počet účastníků a připravit testovací prostředí dle standardizovaných doporučení [8] a [9]. Při závěrečném vyhodnocení výsledků objektivních i subjektivních metod je vhodné vypočítat i korelační koeficient ke zjištění míry korelace mezi jednotlivými metrikami.

### 3.1 Objektivní metriky hodnocení kvality videa

Objektivní metody pro hodnocení kvality videosekvencí se dělí v závislosti na použití původního nekomprimovaného videa. Jsou metody, které srovnávají původní nekomprimované video s komprimovaným, ty se označují jako referenční (plně či částečně). Naopak metody, které původní nekomprimované video vůbec nepoužívají a pracují pouze s komprimovanou sekvencí, se nazývají bezreferenční. Pro účely této práce budou uvažovány pouze referenční a částečně referenční metriky.

#### 3.1.1 PSNR

Zkratka této objektivní metody vychází z anglického názvu Peak Signal-to-Noise Ratio [2]. Jedná se o špičkový poměr signálu vůči šumu. Je velice často využívanou plně referenční metodou. Největší výhodou této metody je jednoduchost její implementace spolu s nízkou náročností výpočtu. Mezi její nevýhody však patří fakt, že algoritmus výpočtu této metody vůbec nezohledňuje vlastnosti lidského zraku. Šum nemusí v každé scéně působit rušivě, PSNR jej však vyhodnotí pro každou scénu shodně. Vzorec pro výpočet PSNR je následující [2]:

$$PSNR = 10 \log_{10} \left( \frac{MAX^2}{MSE} \right) [dB], \quad (3.1)$$

kde  $MAX$  znamená maximální hodnotu pixelu v obrázku (nejčastěji 255 pro 8-bitovou

hloubku) a  $MSE$  je hodnota střední kvadratické chyby, jejíž hodnotu získáme ze vztahu [2]:

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2, \quad (3.2)$$

kde  $N$  je hodnota celkového počtu pixelů a  $x_i$  a  $y_i$  jsou hodnoty jednotlivých pixelů z originálního a komprimovaného signálu.

I přes uvedené nedostatky lze PSNR použít jako hrubý prvotní odhad obrazové kvality. Metoda PSNR bývá také díky své jednoduchosti součástí řídicích systémů zpracovávajících obrazový signál, kde poskytuje zpětnou vazbu nastavení kódování. Výsledná hodnota PSNR je udávána v jednotkách dB. Čím vyšší hodnoty PSNR dosahuje, tím větší by měla být kvalita obrazu. Hodnoty typické pro ztrátovou kompresi videa s 8-bitovou hloubkou se pohybují v rozmezí 20 – 40 dB pro nižší datové toky a 30 – 50 dB pro vysoké datové toky [10].

### 3.1.2 SSIM

Označení SSIM je zkratkou anglického názvu Structural Similarity Index neboli index strukturální podobnosti [10]. Jedná se o plně referenční metriku, která byla navržena tak, aby ve srovnání s PSNR lépe reflektovala vlastnosti lidského zraku. Při hodnocení kvality videa pracuje pouze s jasovou složkou obrazu, neboť vychází z předpokladu, že lidský zrak se zaměřuje spíše na celkovou strukturu scény než na jednotlivé pixely.

Výpočet indexu SSIM je rozdělen do tří hlavních částí, kterými jsou srovnání jasu, kontrastu a struktury. Nejprve je nutné vypočítat průměrnou hodnotu jasu  $\mu_x$  pro zdrojový signál  $x_i$  dle rovnice [10]:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i. \quad (3.3)$$

Výslednou hodnotu je nutné odečíst od zdrojového signálu  $x - \mu_x$ . Dále je nutné provést odhad kontrastu signálu, prostřednictvím směrodatné odchylky  $\sigma_x$ . Ten je určen rovnicí [10]:

$$\sigma_x = \left( \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}}, \quad (3.4)$$

V dalším kroku jsou oba signály normalizovány, když jsou poděleny směrodatnou odchylkou  $(x - \mu_x)/\sigma_x$  a  $(y - \mu_y)/\sigma_y$ . Následně jsou všechny tři části sloučeny do jedné funkce [10]:

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)), \quad (3.5)$$

kde  $l$  je funkce pro porovnání jasu,  $c$  je funkce pro porovnání kontrastu a  $s$  je funkce srovnávající obrazovou strukturu. Konkrétní porovnání jasové složky je provedeno pomocí následujícího vztahu [10]:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (3.6)$$

Obdobně pro srovnání kontrastu platí vztah [10]:

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (3.7)$$

kde  $C_1$  a  $C_2$  jsou konstanty, které se do vztahu přidávají pro zachování stability rovnice a platí pro ně následující vztahy [10]:

$$C_1 = (K_1 L^2), \quad (3.8)$$

kdy hodnotu konstanty  $K_1$  zvolíme tak, že  $K_1 \ll 1$ .

$$C_2 = (K_2 L^2), \quad (3.9)$$

kde hodnotu konstanty zvolíme opět  $K_2 \ll 1$ , avšak různou od  $K_1$ .

Následně je nutné ještě dopočítat hodnotu korelace obou signálů dle vztahu [10]:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (3.10)$$

Nyní je možné vypočítat vzájemné srovnání obrazové struktury dle rovnice [10]:

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x + \sigma_y + C_3}, \quad (3.11)$$

kde hodnotu konstanty  $C_3$  získáme ze výpočtem  $C_3 = C_2 / 2$ .

Nyní, po sloučení všech tří vypočítaných funkcí, je možné tyto dosadit do výsledného vztahu pro index SSIM daného rovnicí [10]:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3.12)$$

Výsledná hodnota indexu SSIM nabývá hodnot od -1 do 1, přičemž hodnota 1 vyjadřuje shodný obraz na vstupu i na výstupu [2].

### 3.1.3 VQM

Video Quality Model (VQM) je plně či částečně referenční metrika hodnocení kvality videosekvencí. Ve srovnání s výše uvedenými objektivními metrikami se jedná o komplexní, výpočetně náročnou metodu, pomocí které však získáme výsledky hodnocení obrazové kvality, které dosahují vyšší míry korelace v porovnání s lidským zrakem. Tato metrika byla navržena tak, aby v co nejvyšší možné míře zohledňovala model lidského vidění, tzv. Human Visual System (HVS). Obsahuje celkem 9 odlišných modelů výpočtu, avšak standardizovány ITU-T byly pouze 3, a to v dokumentech ITU-



T J.144 [11], ITU-T J.244 [12] a ITU-T J.249 [13]. Jedná se o The NTIA General Model, Fast Low Bandwidth Model a PSNR Model.

NTIA General Model je základním, částečně referenčním modelem této metriky, který je optimalizován tak, aby jeho výsledky co nejlépe korelovaly s výsledky subjektivního hodnocení kvality videosekvencí. Tento model byl schválen skupinou expertů VQEG (Video Quality Experts Group) a následně standardizován ITU a publikován v dokumentu ITU-T J.144 [11]. V tomto standardu je zahrnuto taktéž několik kalibračních modelů, které je možné provést před samotným výpočtem VQM. Kalibrace v tomto případě slouží k prostorovému a časovému zarovnání původní a výstupní sekvence. Proveden je taktéž výpočet celkového posunu jasové složky obrazu a určení relevantních částí obrazu. To znamená, že zpracovávány tak nemusí být například okrajové části obrazu, které nenesou žádnou obrazovou informaci. Kalibraci je vhodné použít zejména při přenosu videa, neboť v tomto případě může být obraz doplněn o oblast, která nese užitečnou obrazovou informaci, což by mohlo negativně ovlivnit výsledky měření. Pro účely této práce budou sekvence pouze komprimovány, a nebude proto nutné kalibraci provádět.

Výpočet hodnoty VQM pro NTIA General Model je založen na lineární kombinaci celkem sedmi nezávislých parametrů, které byly navrženy tak, aby co nejvíce reflektovaly zpracování obrazu lidským vizuálním systémem HVS. Z uvedených sedmi parametrů jsou čtyři založeny na změně jasové složky v obraze, dva na kombinaci chrominančních složek a poslední představuje měření kontrastu a pohybu prostřednictvím jasové složky obrazu. Vzorec pro lineární kombinaci uvedených parametrů má tuto podobu [14]:

$$\begin{aligned} VQM = & 0,297 * si\_loss \\ & +0,5969 * hv\_loss \\ & +0,2483 * hv\_gain \\ & +0,0192 * chroma\_spread \\ & -2,3416 * si\_gain \\ & +0,0431 * ct\_ati\_gain \\ & +0,0076 * chroma\_extreme. \end{aligned} \quad (3.13)$$

Parametr „*si\_loss*“ zaznamenává ztrátu prostorové informace, například rozmazání. K výpočtu používá tento parametr 13-pixelový filtr prostorové informace s označením SI13.

Parametr „*hv\_loss*“ detekuje posun hran z horizontální a vertikální orientace pro případ, kdy by z důvodu rozmazání došlo k většímu poškození horizontálních a vertikálních hran jasové složky oproti šikmým. Tento parametr používá horizontálně a vertikálně filtrovaný obraz na výstupu z filtru SI13.

Parametr „*hv\_gain*“ zjišťuje změnu hran ze šikmých na horizontální či vertikální, což se stává zpravidla v případě, kdy výstupní video obsahuje tzv. blokové artefakty.

Parametr „*chroma\_spread*“ sleduje změny v rozložení chrominančních složek obrazového signálu. Jinak řečeno, detekuje změnu barev výstupní sekvence oproti originální.

Parametr „*si\_gain*“ zaznamenává zlepšení kvality výstupní sekvence, například pomocí

zostření hran obrazu či zvýšení kontrastu.

Parametr „*ct\_ati\_gain*“ porovnává hodnoty prostorové a časové aktivity obrazu, neboť při vysokých hodnotách nebudou šum a chybné bloky v obraze tolik patrné. Důležitá je v tomto případě také hodnota kontrastu obrazu.

Parametr „*chroma\_extreme*“ hlídá vážné, avšak lokalizované chyby v barevnosti, které vznikají zejména v obrazových přenosových systémech a jsou pro diváka velmi rušivé.

Výsledné hodnoty VQM se pohybují od 0 do 1 s tím, že 0 reprezentuje shodný obraz na vstupu i výstupu a 1 reprezentuje velmi markantní degradaci výstupního obrazu oproti vstupnímu.

### 3.1.4 VIF

Metoda Visual Information Fieldity (VIF) je podobně jako VQM komplexní metodou objektivního hodnocení obrazové kvality s vysokou mírou korelace se subjektivním hodnocením. Také v tomto případě se jedná o složitý a výpočetně náročný algoritmus. Výsledná hodnota této metriky je odvozena od srovnání vzájemné informace na vstupu a výstupu modelu lidského vizuálního systému HVS. Tato metoda využívá komplikované stochastické modely HVS spolu se zdrojovým a komprimovaným signálem. Na zdrojový i komprimovaný obraz je aplikována zvláštní sada filtrů, které se snaží co nejvíce přiblížit zpracování signálu zrakovým centrem mozku. Tyto filtry byly vybrány na základě rozsáhlého výzkumu subjektivního porovnání kvality videosekvencí [15].

Samotný výpočet je prováděn ve frekvenční oblasti, do které je signál převeden prostřednictvím diskrétní vlnkové transformace DWT (Discrete Wavelet Transform). Hlavními modely se kterými VIFP pracuje jsou Gaussovo rozložení spolu s modelem aditivního šumu [15]. Měřítkem kvality je pak informace, která je společná pro vstupní i výstupní signál. Jako ukazatel slouží vypočítané množství informací, které je možné získat ze signálu komprimovaného oproti referenčnímu. Výsledný index VIF nabývá hodnot od 0 do 1, přičemž hodnota 1 vyjadřuje shodný obraz na vstupu i výstupu.

Pro účely této práce bude použita zjednodušená verze této metriky s označením VIFP (Visual Information Fieldity in Pixel Domain), která srovnává věrnost obrazové informace na úrovni jednotlivých pixelů.

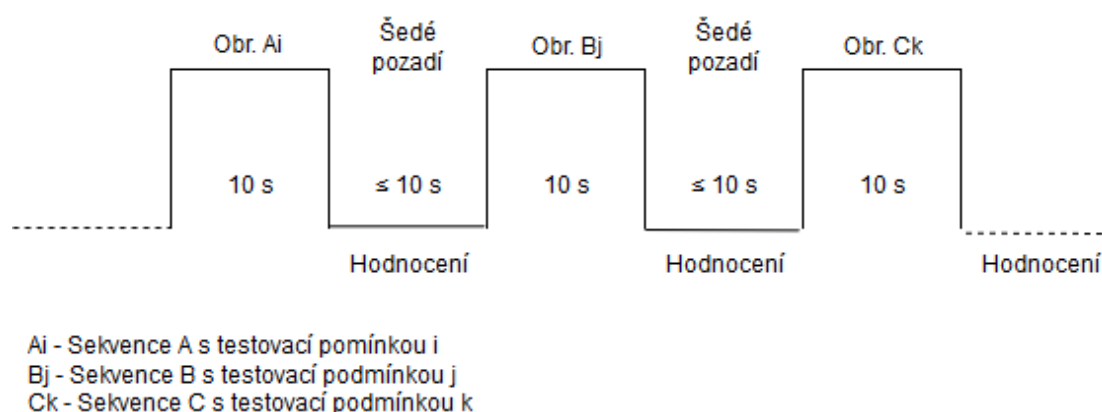
## 3.2 Subjektivní metriky hodnocení kvality videa

Princip subjektivních metod hodnocení kvality videosekvencí spočívá v představení krátkých videosekvencí vybranému reprezentativnímu počtu pozorovatelů, kteří posoudí kvalitu komprimovaných videí. Následně je vhodně zvolenou statistickou metodou určen výsledný index kvality. Obdobně jako u objektivních metrik, můžeme i tyto metody rozdělit podle toho, zda bude pozorovateli poskytnut pouze jeden podnět (*single stimulus*), dva podněty (*double stimulus*), případně více podnětů (*multi stimulus*). U subjektivního hodnocení však hraje dále roli nejen zvolená videosekvence, ale určitý vliv má také okolní prostředí, které může pozorovatele ovlivnit. Z toho důvodu vypracoval institut ITU-T standardizované postupy pro použití jednotlivých metrik. Ty jsou uvedeny v dokumentech ITU-T P.910 [8] a ITU-T BT.500-13 [9].

### 3.2.1 ACR

Subjektivní metoda s označením ACR (Absolute Category Rating) je typickým zástupcem metody s jedním podnětem (single stimulus). U této metody je pozorovateli přehráván náhodný sled krátkých videosekvencí (typicky 8 – 10-vteřinových), mezi kterými je vždy zařazena pauza o maximální délce deseti vteřin. Během této pauzy pozorovatel subjektivně hodnotí kvalitu sekvence, kterou právě viděl, bez ohledu na možnost porovnání s originální nekomprimovanou sekvencí. Subjektivní hodnocení kvality pozorovatel zaznamená do dotazníku a ohodnotí známkou od 1 (nejhorší kvalita) do 5 (nejlepší kvalita). Pokud je z nějakého důvodu požadováno podrobnější rozlišení, připouští doporučení ITU-T použití 9 stupňového (0 až 9) dotazníku, který mezi již existující stupně přidává jeden mezistupeň.

Schéma testování sekvencí metodou ACR je uvedeno na následujícím Obrázek 3-1. Vždy po přehrání testované sekvence by mělo být na obrazovce zobrazeno šedé pozadí, a to po dobu nejdéle 10 vteřin, během které pozorovatel provede hodnocení kvality. Hodnocení může být zaznamenáno do dotazníku, případně je možné hodnocení zaznamenat interaktivně na připojeném PC. Po hodnocení následuje jiná testovací sekvence.



Obrázek 3-1 Schéma postupu testování metodou ACR

Tabulka 3-1 Kategorie ACR (podle [9])

Známka	Hodnocení
5	Výborný
4	Dobrý
3	Příjemný
2	Ucházející
1	Nepříjemný

Výsledný index kvality získáme aritmetickým průměrem všech výsledků pozorování. Tento výsledek se nazývá MOS (Mean Opinion Score). Do výsledného indexu kvality MOS by měla být připočtena ještě nejistota měření, kterou určíme ze směrodatné odchylky a spolehlivosti 95 % dle následujícího vztahu [9]:

$$u_c = 1,96 \frac{s}{\sqrt{N}}, \quad (3.14)$$

kde  $N$  označuje počet zúčastněných pozorovatelů a  $s$  značí výběrovou směrodatnou odchylku, kterou vypočítáme ze souboru udělených známek následovně [9]:

$$s = \sqrt{\sum_{i=1}^N \frac{(\bar{v} - v)^2}{(N-1)}}, \quad (3.15)$$

kde  $\bar{v}$  odpovídá průměrné známce MOS skóre a  $v$  je skóre, které udělili jednotliví pozorovatelé.

### 3.2.2 DSCQS

Subjektivní metoda hodnocení s označením DSCQS (Double Stimulus Continous Quality Scale) zahrnuje taktéž dva podněty (double stimulus). V tomto případě je však pozorovateli prezentována sekvence referenčního a kódovaného videa v náhodném pořadí, přičemž je požádán, aby zhodnotil kvalitu obou sekvencí [9].

Tato metoda má ještě dvě varianty provedení. První varianty se účastní jediný pozorovatel, který má možnost volně přepínat mezi dvěma sekvencemi dlouhými opět 10 sekund, dokud se sám nerozhodne, jak kterou sekvenci ohodnotit. U druhé varianty hodnotí testované sekvence více pozorovatelů najednou, přičemž jsou tyto sekvence spuštěny opakovaně. Hodnocení probíhá až během posledního opakování sekvence [9].

Pozorovatel má pro hodnocení k dispozici spojitou stupnici s pětistupňovou škálou, na kterou postupně vyznačí hodnocení všech shlédnutých sekvencí. Dotazník k hodnocení dle doporučení uvedeného v [9] ilustruje Obrázek 3-2.

	01	02	03	04				
	A	B	A	B	A	B	A	B
Výborný								
Dobrý								
Příjemný								
Ucházející								
Nepříjemný								

Obrázek 3-2 Hodnotící formulář DSCQS (na základě [1])

Při zpracování výsledků je nutné nejprve určit rozdíl mezi testovanou a referenční sekvencí a následně se vypočítá střední hodnota pro každého pozorovatele. Pro každou testovací sekvenci je nutné vypočítat střední hodnotu  $\bar{u}_{jkr}$ , dále standardní odchylku  $S_{jkr}$  a koeficient  $\beta_{2jkr}$  podle následující rovnice [9]:

$$\beta_{jkr} = \frac{m_4}{(m_2)^2}, \quad (3.16)$$

kdy koeficient  $m_x$  získáme ze vztahu [9]:

$$m_x = \frac{\sum_{i=1}^N (u_{ijkr} - \bar{u}_{jkr})^x}{N} \quad (3.17)$$

## 4 POSTUPY PRO KOMPRIMACI VIDEOSEKVENCÍ

### 4.1 Referenční implementace VP9

Referenční implementace kodeku VP9, která je součástí multimediální knihovny *libvpx*, byla opatřena z oficiálního webu projektu WebM [16]. Jelikož se jedná o volnou (open source) licenci, je možné na webu stáhnout zdrojový kód určený pro kompilaci. Aktuální verze multimediální knihovny *libvpx* je označena 1.6.1 [16].

#### 4.1.1 Kompilace programu

Samotný proces kompilace programu je poněkud složitější, neboť software je zaměřený na uživatele systému Linux. Aby bylo možné zkompilevat program i v systému Windows, je nutné si opatřit například emulátor příkazové řádky linuxového prostředí (MSYS + MiniGW [17]). Na webu projektu WebM [16] je uvedeno několik postupů, jak s pomocí linuxového emulátoru a programu MS Visual Studio zkompilevat výsledný program.

Zkompilováním vznikl program s názvem *vpxenc.exe*, sloužící pro kódování videa, a dále soubor s názvem *vpxdec.exe*, který slouží naopak pro jeho dekódování.

#### 4.1.2 Příkazy

Program se ovládá prostřednictvím příkazového řádku a požadované nastavení se přidává jako parametr programu s následující syntaxí:

```
-parametr hodnota  
--parametr=hodnota
```

Následující Tabulka 4-1 uvádí stručný přehled základních parametrů sloužících pro nastavení kodéru VP9.

Tabulka 4-1 Příkazy syntaxe VP9 (na základě [16])

Výstup	Název a cesta k výstupnímu souboru
codec	Volba kodeku, pro VP9 vpxenc-vp9
limit	Po N-tém snímku se ukončí kódování
skip	Přeskočit prvních N-snímku
profile	Volba profilu
ewidth w-	Nastavení šířky videa
height h-	Nastavení výšky videa
fps	Počet snímků za vteřinu
target_bitrate	Cílová přenosová rychlost
min_q	Minimální hodnota kvantizéru
max_q	Maximální hodnota kvantizéru
i420	Vstupní video je ve formátu i420
psnr	Po dokončení zobrazí PSNR
passes	Jedno nebo dvoupásmové kódování

Konkrétní zadání syntaxe příkazové řádky pak vypadá takto:

```
vpxenc.exe input_videoUHD_30fps.yuv -o output_video_vp9.webm -
codec=vp9 -w 3840 -h 2160 -fps=30 -p420 -target-bitrate=12000 -minq=0
-max-q=60
```

V syntaxi je nejprve zadán zdrojový soubor určený ke kompresi, jelikož formát „yuv“ obsahuje pouze obrazová data (nejsou obsaženy základní informace o sekvenci jako v multimediálním kontejneru, např. avi). Dále je nutné specifikovat rozlišení zdrojové sekvence, snímkovou rychlost a barevný model (v tomto případě YUV 4 : 2 : 0p). Dále je specifikován cílový bitový tok sekvence a rozptýl nastavení kvantizéru.

## 4.2 Referenční implementace HEVC (H.265)

Vzhledem k faktu, že se jedná o profesionální licencovaný software, je možné pro zkušební účely stáhnout referenční implementaci HEVC s označením HM (HEVC Test Model [18]. v aktuální verzi HM-16.8). Zkompilování programu je ve srovnání s VP9 jednodušší díky nativní podpoře systému Windows. Pro účely zkompilování můžeme zvolit například předem připravený projekt v prostředí programu MS Visual Studio.

### 4.2.1 Ovládání kodeku

Po zkompilování jsme získali programy *TAppEncoder.exe* a *TAppDecoder.exe* pro kódování a dekódování videa. Ovládání je podobné jako u VP9. Kodér spustíme prostřednictvím příkazové řádky. Příkazy pro ovládání kodeku můžeme dávat pomocí parametrické syntaxe stejně jako v případě VP9, nicméně HEVC podporuje funkci čtení příkazů z konfiguračního souboru (*config.cfg*). To je výhodné zejména při kódování většího množství sekvencí, kdy mohou být parametry pevně definovány v konfiguračním souboru.

Obecná syntaxe ovládání dekodéru vypadá následovně:

```
TAppEncoder.exe [-h] [-c config.cfg] [--parametr=hodnota]
```

Zadáním parametru [-h] se na obrazovce objeví výpis použitých parametrů. Při zadání parametru [-c název.cfg] je kodér informován o načtení parametrů z konfiguračního souboru včetně jejich názvu a cesty. Pokud chceme vkládat parametry přímo, pak použijeme výše uvedenou syntaxi, prakticky totožnou s VP9. Následující Tabulka 4-2 obsahuje přehled základních parametrů pro ovládání kodéru HEVC. Vzhledem k tomu, že kodér HEVC disponuje opravdu velkým množstvím příkazů, ilustruje tabulka pouze několik základních typů.

Tabulka 4-2 Příkazy syntaxe HEVC (na základě [19])

Výstup	Název a cesta k výstupnímu souboru
InputFile	Vstupní sekvence
BitstreamFile	Výstupní kódované video
ReconFile	Výstupní rekonstruovaný soubor po kompresi
SourceWidth	Šířka vstupní sekvence
SourceHeight	Výška vstupní sekvence
OutputBitDepth	Bitová hloubka rekonstruovaného videa
FrameRate	Počet snímků za vteřinu u vstupního videa
FrameSkip	Počet snímků, které kodér na začátku preskočí
FramesToBeEncoded	Počet snímků, které mají být kódovány
Profile	Volba profilu kodéru
Level	Volba úrovně kodéru
MaxCUWidth	Definuje maximální velikost šířky makrobloků
MaxCUHeight	Definuje max. velikost výšky makrobloků
MaxCUSize	Maximální velikost makrobloku
QP	Definuje hodnotu kvantizačního parametru
LoopFilterDisable	Povolí nebo zakáže deblock loop filtr

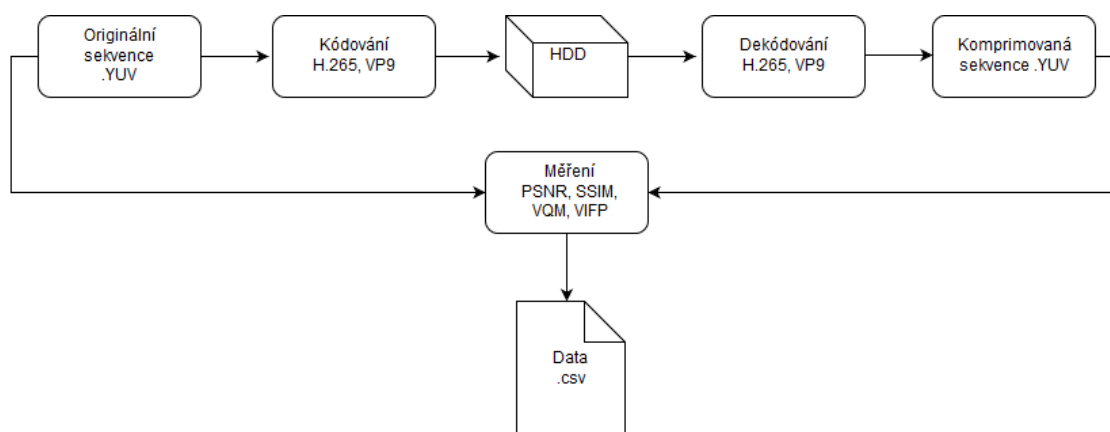


## 5 SROVNÁNÍ KOMPRESNÍ ÚČINNOSTI JEDNOTLIVÝCH STANDARDŮ

V následující části této práce bude podrobně popsáno provedené porovnání výkonnosti kódovacích standardů VP9 a H.265, a to prostřednictvím několika krátkých testovacích sekvencí ve vysokém (4K) a nízkém rozlišení, které byly komprimovány oběma kodeky s nastavením různých bitových rychlostí. Kvalita komprimovaných sekvencí byla následně vyhodnocena výše popsánymi objektivními metrikami. Mimo vyhodnocení kvality sekvencí objektivními metrikami byly také realizovány subjektivní testy, a to – dle doporučení ITU-T – výše popsanou metodou ACR (Absolute Category Rating). V závěru této části bude provedeno také srovnání všech dosažených výsledků hodnocení kvality videosekvencí a taktéž připraven návrh laboratorní úlohy.

### 5.1 Metoda srovnání

Porovnání účinnosti jednotlivých kodeků bylo provedeno prostřednictvím kódování krátké testovací videosekvence. Základem je tedy nekomprimovaná videosekvence ve formátu „yuv“. Pro danou videosekvenci je nutné nastavit potřebné parametry a všechny sekvence postupně komprimovat oběma kodeky. Po zkomprimování byly sekvence uloženy na hard disk (HDD) v příslušných multimediálních kontejnerech, s těmito ale nelze provést referenční srovnávací hodnocení. Z toho důvodu je nutné zkomprimované sekvence převést zpět do nekomprimovaného formátu „yuv“. Poté je již možné provést referenční vyhodnocení kvality jednotlivých komprimovaných sekvencí. Pro každou zkomprimovanou sekvenci tak byly získány hodnoty PSNR, SSIM, VQM a VIFP a tyto byly spolu s bitovou rychlostí vyneseny do jednotlivých výsledných grafů.



Obrázek 5-1 Diagram postupu srovnání sekvencí

### 5.2 Použitý software

Vzhledem k vysoké hardwarové a časové náročnosti používání referenčních implementací jednotlivých kódovacích standardů byl ke kompresi testovacích

videosekvencí použit volně dostupný software FFmpeg ve verzi *20160330-git-be746ae*, dostupný z [20]. Tento software obsahuje knihovny obou uvažovaných kódovacích standardů. Je optimalizován, takže zvládá komprimovat sekvence mnohem rychleji, než je tomu u referenčních implementací. Ovládá se však naprosto stejně jako referenční implementace jednotlivých kodeků, příkazy jsou shodně vkládány do příkazové řádky. Umožňuje také využití dávkového souboru, což velmi usnadňuje práci, neboť není nutné všechny potřebné údaje psát opakovaně do příkazové řádky pro každou sekvensi zvlášť. Díky tomu bylo možné zpracovat velký počet sekvencí s proměnnými parametry.

Pro komplexní výpočet objektivních měřících technik zkomprimovaných videosekvencí byl použit software VQMT (Video Quality Measurement Tool) ve verzi 1.1, který má v sobě implementovány námi požadované objektivní metriky, tedy kromě VQM, a který je volně dostupný pro studijní účely z [21]. Také tento software se ovládá pomocí parametrů příkazové řádky, stejně tak je možné si připravit dávkový soubor pro zpracování velkého počtu sekvencí. Výstupem tohoto programu jsou „csv“ soubory pro každou měřenou sekvensi a každou jednotlivou metriku.

Za účelem výpočtu hodnot metriky VQM byl použit software s označením BVQM (Batch Video Quality Metric) ve verzi 2.1, který je pro akademické účely dostupný z [22]. Program se ovládá pomocí grafického rozhraní, samotný výpočet pak probíhá pomocí programu MATLAB. Software umožňuje volbu jednoho z 9 výpočetních modelů, dále 8 možných nastavení kalibrace nebo také manuální vložení kalibračních dat. Pro účely této práce byl použit základní, standardizovaný výpočetní model NTIA General Model. Kalibrace v tomto případě použita nebyla, neboť při kompresi zpravidla nedochází k posunu či doplnění částí obrazu. Výstupem z tohoto programu jsou jednak sloupcové grafy zpracované programem MATLAB a dále data v souborech „csv“, která obsahují jednak průměrnou hodnotu VQM pro každou testovanou sekvensi a také hodnoty jednotlivých parametrů, ze kterých se výsledný VQM index skládá.

## 5.3 Videosekvence

Za účelem srovnání účinnosti výše uvedených kódovacích standardů byly použity celkem 4 desetiřádkové nekomprimované videosekvence v rozlišení 4K (3860 x 2160 pixelů), formátu „yuv“ 4 : 2 : 0, bitové hloubce 8 a snímkové rychlosti 30 fps. Každá sekvence tedy obsahuje 300 snímků. Sekvence byly zvoleny tak, aby obsahovaly statické i dynamické scény. Testovací sekvence byly získány od Shanghai Jiao Tong University (SJTU) [23], která disponuje databází nekomprimovaných 4K videosekvencí, které jsou volně dostupné pro studijní účely.

Tyto sekvence byly následně zkomprimovány oběma kodeky s nastavením pěti různých bitových rychlostí: 2, 6, 12, 20 a 40 Mbit/s. Bitová rychlost byl základní proměnný parametr pro všechny použité sekvence, ostatní nastavení obou kodérů bylo ponecháno na výchozím nastavení. U H.265 byl zvolen základní profil s označením „main“ a u VP9 byl specifikován rozsah hodnot kvantizéru na rozptýl hodnot 0 až 60, protože dle doporučení vývojářů [24] a dostupných studií [25] při tomto nastavení dosahuje kodér nejlepších výsledků. Rozptýl bitových rychlostí byl záměrně zvolen tak, aby byly při nejnižší hodnotě patrné degradace obrazu až po nejvyšší, u které by rozdíl

mezi kódovanou a původní sekvencí by neměl být postřehnutelný.

Obsah příkazové řádky pro kódování sekvence Fountains v UHD rozlišení s cílovou bitovou rychlostí 2 Mbit/s, pomocí kodeků H.265:

```
ffmpeg -f rawvideo -vcodec rawvideo -s 3840x2160 -r30 -pix_fmt yuv420p  
-i Fountains.yuv -c:v libx265 -profile: v main -b: v 2M output.mkv
```

A obdobně pro kompresi pomocí kodeku VP9:

```
ffmpeg -f rawvideo -vcodec rawvideo -s 3840x2160 -r30 -pix_fmt yuv420p  
-i Fountains.yuv -c:v libvpx-vp9 -qmin 0 -qmax 60 -b: v 2M output.webm
```

Z obsahu syntaxe příkazové řádky je patrné, že po provedení komprese byly uvedené sekvence komprimovány v příslušných multimediálních kontejnerech („mkv“ pro H.265 a „webm“ pro VP9). Dle schématu uvedeném na Obrázek 5-1 bylo nutné v následujícím kroku provést dekódování všech sekvencí zpět do formátu „yuv“.

Po zkomprimování všech sekvencí ve vysokém rozlišení pomocí obou kodérů byl proveden bezztrátový převod všech originálních videosekvencí do nízkého rozlišení o velikosti 854 x 480 pixelů a poté opět provedena komprese oběma standardy, avšak s nastavením odlišných bitových rychlostí o velikosti 250, 750, 1500 a 4000 kbit/s.

### 5.3.1 Přehled videosekvencí

Pro ilustraci uvádím na tomto místě přehled testovaných videosekvencí. Po stručném popisu jednotlivých videí následuje Tabulka 5-1, ve které jsou zobrazeny i náhledy pro každé jednotlivé video.

- |                           |  |
|---------------------------|--|
| <b>a) Tall Buildings:</b> | Statická denní scéna zachycující mrakodrapy a pohybující se dopravu v pozadí.  |
| <b>b) Campfire Party:</b> | Dynamická noční scéna s plameny a pohybujícími se postavami v pozadí.  |
| <b>c) Fountains:</b>      | Dynamická denní scéna se stříkající fontánou, čeřící se vodní hladinou pod dopadajícími kapkami vody a statickými budovami v pozadí. |
| <b>d) Wood:</b>           | Statická denní scéna v lese, kamera se pomalu pohybuje do strany a sluneční paprsky prosvítají kolem stromů.                         |

Tabulka 5-1 Přehled jednotlivých videosekvencí



a) Tall Buildings



b) Campfire Party



c) Fountains



d) Wood

### 5.3.2 Informace o prostorové a časové aktivitě (SI a TI)

Pro hlubší popis a analýzu zkoumaných sekvencí je vhodné provést měření prostorových parametrů TI (Temporal Information) a časových parametrů SI (Spatial Information). Hodnoty SI a TI slouží i k výběru testovacích sekvencí, aby k testování byly vybrány sekvence s různorodou velikostí časové a prostorové aktivity.

Měřením prostorových parametrů SI získáme informaci o maximální hodnotě míry prostorových detailů dané sekvence. Vysokou hodnotu mají zpravidla velmi komplexní scény. Výpočet hodnoty prostorových parametrů je založen na Sobelově filtru, kterým je filtrována jasová složka jednotlivých snímků v čase a následně je vypočtena směrodatná odchylka každého pixelu ve srovnání s původním snímkem. Postup výpočtu SI parametru udává následující rovnice [8].

$$SI = \max_{time} \{std_{space}[Sobel(F_n)]\}, \quad (5.1)$$

kde  $F_n$  označuje snímek v čase  $n$  po aplikaci Sobelova filtru,  $std_{space}$  značí směrodatnou odchylku snímku a  $\max_{time}$  označuje maximální hodnotu v daném čase.

Informace o časové aktivitě TI udává hodnotu maximální míry časové změny videosekvence. Obecně lze říci, že čím dynamičtější je scéna, tím vyšší by měla být hodnota časové aktivity. Výpočet je založen na měření hodnoty rozdílu  $M_n(i,j)$  mezi jasovými pixely na stejných pozicích dvou po sobě jdoucích snímků, což udává rovnice [8]:

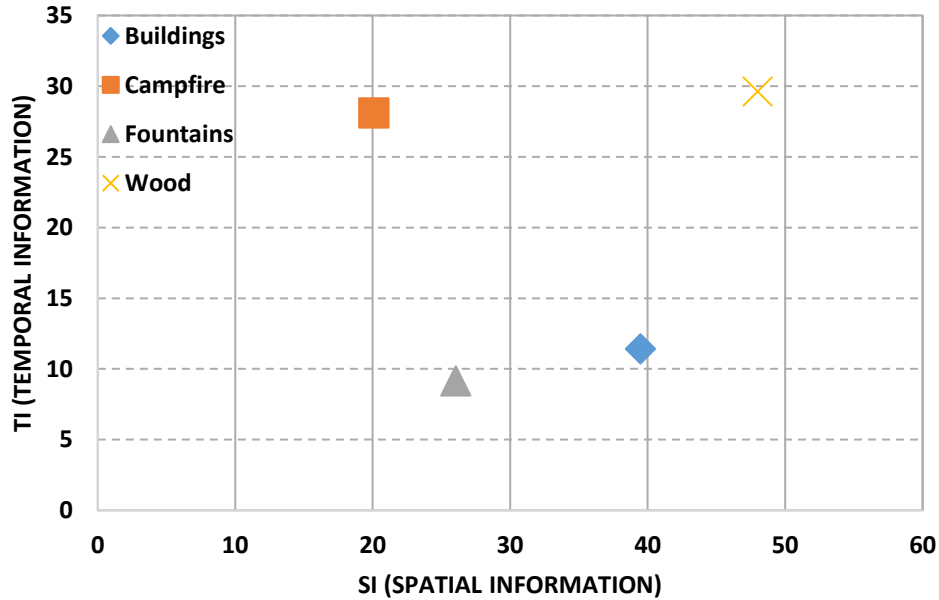
$$M_n(i,j) = F_n(i,j) - F_{n-1}(i,j), \quad (5.2)$$

kde  $F_n(i,j)$  je pixel na  $i$ -tém řádku a  $j$ -tém sloupci snímku v čase.

Výsledná hodnota časové aktivity TI je pak vypočtena obdobně jako SI [8]:

$$TI = \max_{time} \{std_{space}[M_n(i,j)]\}, \quad (5.3)$$

kde  $std_{space}$  udává směrodatnou odchylku rozdílu jednotlivých pixelů mezi snímky a  $\max_{time}$  je maximální hodnota v uvažovaném čase.



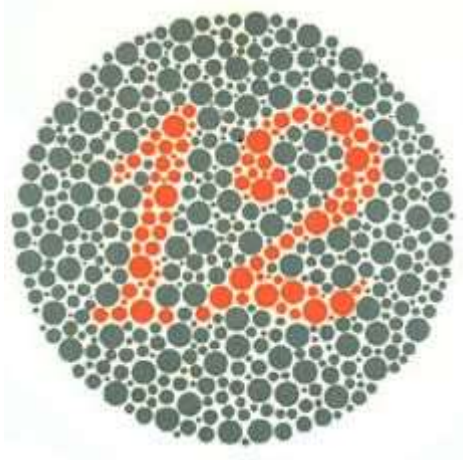
Obrázek 5-2 Informace o prostorové a časové aktivitě

Výsledné hodnoty pro všechny zdrojové sekvence v rozlišení UHD jsou zobrazeny na Obrázek 5-2. Jak bylo již uvedeno dříve, zaznamenány byly maximální hodnoty SI a TI ze všech snímků pro každou testovanou sekvenci. Z grafu je patrné, že pro účely testování byly videosekvence vhodně zvoleny, neboť hodnoty parametrů SI a TI jsou rozprostřeny v grafu a jedná se tak o žádoucí kombinaci statických a dynamických scén, stejně jako komplexních i jednodušších. Tato různorodost zajistí požadavek na různý kompresní poměr u každé jednotlivé sekvence, což je určující pro nastavení různých bitových rychlostí, které jsou v tomto případě základním proměnným parametrem. Samotný výpočet SI a TI parametrů byl proveden prostřednictvím skriptu v programu MATLAB, a to dle doporučení ITU-T P.910 [8].

## 5.4 Subjektivní testy

Za účelem ověření výsledků získaných použitím výše uvedených objektivních metrik, bylo provedeno také subjektivní testování kvality komprimovaných videosekvencí. Subjektivní testy byly provedeny metodou ACR, a to přesně dle doporučení ITU BT.500-13 [9] a ITU-T P.910 [8].

Subjektivního testování se zúčastnilo celkem 20 respondentů o průměrném věku 28,1 let. Minimální počet respondentů pro relevantní výsledky je dle doporučení ITU-T stanoven na 15 respondentů. Zásadním požadavkem pro výběr respondentů bylo, aby se nejednalo o profesionály v oblasti hodnocení videosekvencí. Vzorek respondentů byl tedy vybrán náhodně, část tvořili studenti VUT, další část lidé mimo univerzitu. Nejmladšímu z účastníků bylo 19 let, nejstaršímu 61. Dalším požadavkem bylo vyloučení respondentů s těžkým zrakovým postižením, případně ztrátou barvocitu. Každý účastník byl tedy před samotným zahájením testování sekvencí otestován na barvoslepost prostřednictvím Ishihariho barvocitových tabulek [26] na Obrázek 5-3. Tyto tabulky obsahují množství bodů různých barev a různého jasu, přičemž barevné body vytváří obrazec určité číslice, písmena, či geometrického tvaru. Aby respondent tímto testem prošel, musí úspěšně identifikovat každý obrazec. Osoba s porušeným barvocitovým vnímáním není schopna některé znaky identifikovat.



Obrázek 5-3 Ishihariho pseudoizochromatické tabulky (převzato z [26])

Všichni respondenti tímto testem úspěšně prošli a žádný tedy nemusel být z testu vyřazen.

Z důvodu zajištění reprodukovatelnosti tohoto měření byly subjektivní testy uspořádány v laboratorních podmínkách, konkrétně v laboratoři digitální televizní a rozhlasové techniky na Ústavu radioelektroniky VUT. Zde bylo vytvořeno testovací prostředí podle doporučení ITU-T BT.500-13 [9]. Pro účely testování bylo v místnosti připraveno kontrolované ambientní osvětlení, mimo to byla místnost během testování zcela zatemněna. Testovací pracoviště pro hodnocení sekvencí v UHD rozlišení tvořil televizor Samsung UE50JU6900 s úhlopříčkou 127 cm s podporou rozlišení UHD (4K). Vzdálenost pozorovatele od obrazovky je při tomto rozlišení doporučena [9] tak, aby tvořila 1,6 násobek výšky obrazovky, což v našem případě znamenalo vzdálenost 1,5 metru. Pro testování sekvencí v nízkém rozlišení bylo připraveno vedlejší pracoviště s televizorem Samsung T24E390EW, což je zařízení vybavené úhlopříčkou o velikosti 61 cm a rozlišením obrazovky Full HD (1920x1080). Kvůli zajištění optimálních pozorovacích úhlů a podmínek pro každého jednotlivého pozorovatele bylo rozhodnuto, že každého testování se zúčastní maximálně dva pozorovatelé současně, kteří provedou nejprve hodnocení sekvencí ve vysokém rozlišení a následně hodnocení sekvencí s nízkým rozlišením na vedlejším pracovišti.

Každý pár pozorovatelů byl tedy nejprve požádán o provedení testu za účelem zjištění barvocitových defektů zraku a poté jim bylo představeno samotné testovací pracoviště. Každý pozorovatel následně obdržel tištěný formulář, do kterého bylo nutné zaznamenat subjektivní hodnocení kvality každé videosekvence, a to na stupnici od 1 (nejhorší) do 5 (nejlepší). Pro tyto účely bylo předem připraveno několik náhodných seznamů videosekvencí (playlist). V každém bylo náhodné pořadí testovaných sekvencí, takže každý pár pozorovatelů viděl testované sekvence v jiném pořadí. Náhodnost přehrávaných sekvencí pro každý test je vhodné použít z toho důvodu, aby výsledky nemohly být ovlivněny vzájemným sdělováním dojmů účastníků testování. Každý účastník byl také požádán, aby kvalitu hodnotil pouze sám za sebe.

Dle standardizovaných předpisů pro testování metodou ACR [8] byly následně pozorovatelům přehrány náhodně testované videosekvence a mezi každou sekvencí byla vložena pauza o délce 5 vteřin, během které bylo dle doporučení ITU-T [8] na obrazovce promítáno šedé pozadí. Během této pauzy měli pozorovatelé možnost zhodnotit kvalitu předešlé videosekvence. Pro účely testování sekvencí ve vysokém rozlišení bylo divákům přehráno celkem 40 testovacích sekvencí a následně 32 sekvencí o nízkém rozlišení na vedlejším pracovišti. Pro každou sadu sekvencí obdrželi hodnotitelé odlišný hodnotící formulář. Časová náročnost testování byla přibližně 18 minut pro jeden pár pozorovatelů.

Vyhodnocení získaných dat bylo provedeno ručním přepisem jednotlivých hodnot do tabulky programu MS Excel. Ze všech hodnot byl následně vypočten aritmetický průměr MOS a ten následně vnesen do sloupcových grafů, kde byl zohledněn spolu s 95% konfidenčním intervalem, prostřednictvím chybových úseček v grafu.





Obrázek 5-4 Testovací pracoviště

## 5.5 Výsledky

Veškerých výsledků hodnocení kvality sekvencí, které budou uváděny v následující části, bylo dosaženo na stolním PC s procesorem Intel® Core™ i7-4790 @ 3.60 GHz s grafickou kartou Nvidia GeForce GTX 780 a operační pamětí o velikost 16 GB. Časová náročnost jednotlivých výpočtů se výrazně lišila dle komplexnosti jednotlivých metrik. Výpočet indexu PSNR a stejně tak SSIM byl dle předpokladů nejméně časově náročný, doba nutná k výpočtu obou hodnot pro jednu sekvenci byla v řádu několika minut pro sekvence v UHD rozlišení. Výpočet metody VIFP je ve srovnání s předchozími náročnější a komplexnější, jeho výpočet pro jednu testovací sekvenci trval přibližně 30 minut. Nejnáročnější z hlediska výpočetního výkonu je poslední z uvažovaných metod, a to VQM. V tomto případě trval výpočet jedné sady sekvencí v UHD rozlišení (tedy 5 různých bitových rychlostí + referenční) asi 80 minut.

Při vyhodnocení výsledků je nutné vzít v úvahu vzájemnou korelaci jednotlivých objektivních metrik vůči metodě subjektivní. Dle dostupných předpokladů [15] a [25] by nejnižší míry korelace měla dosahovat metrika PSNR. SSIM dle studií dosahuje vyšší míry korelace. Nejvyšší míry korelace by pak měly dosahovat metricky VQM a VIFP. Výsledky jednotlivých objektivních metrik byly vyneseny do spojnicových grafů a zobrazeny v závislosti na dosažené bitové rychlosti.

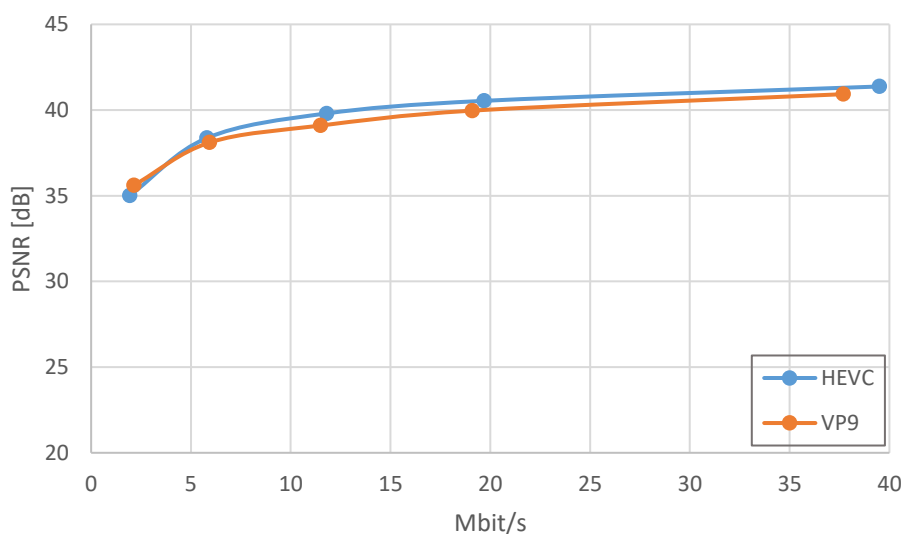
### 5.5.1 Sekvence Tall Buildings

#### Rozlišení UHD (3840 x 2160)

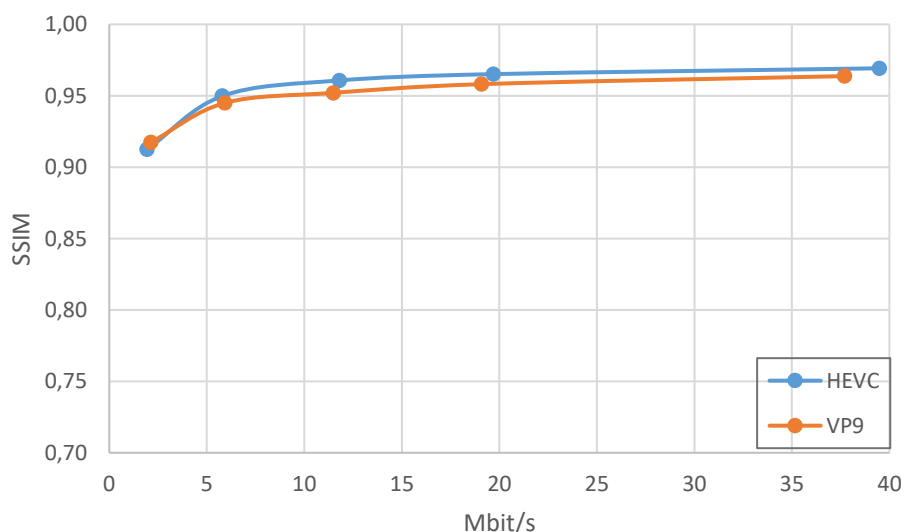
Statická denní scéna (Tabulka 5-1) s nízkou hodnotou prostorové aktivity (TI), zjištěné hodnoty jednotlivých metrik tak reflektují povahu sekvence. Kodeky VP9 a H.265 zde mohou plně využít svých predikčních algoritmů, čímž dochází k výrazné úspoře bitového toku, aniž by byl obraz poznamenán postřehnutelnými degradacemi. Výsledné hodnoty PSNR se u této sekvence pohybovaly v rozmezí 35 – 41 dB. Kodek H.265 dosáhl vysoké hodnoty 40 dB s datovým tokem 12 Mbit/s.



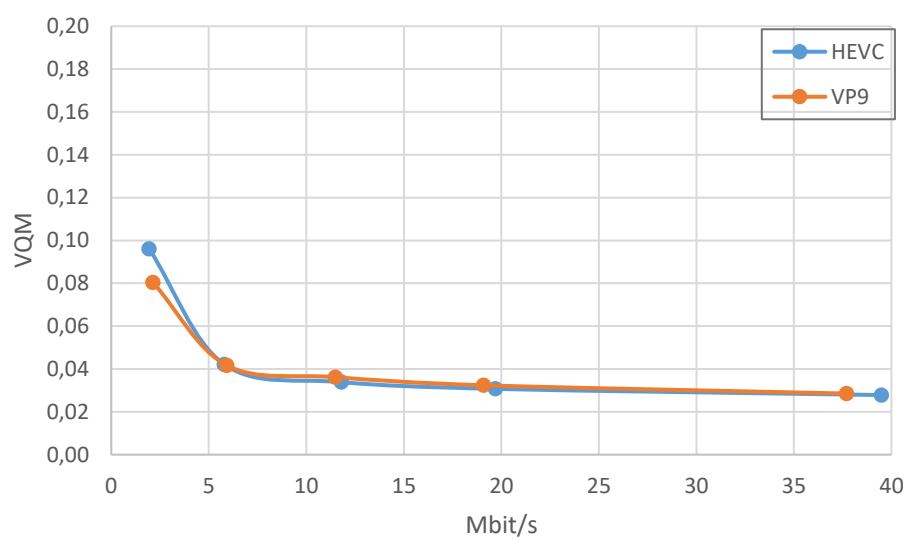
Výsledky SSIM (Obrázek 5-6) se v tomto případě pohybovaly v intervalu od 0,91 do 0,96. Obdobné výsledky jsou patrné v grafech dalších metrik VQM (Obrázek 5-7) a VIFP (Obrázek 5-8). Vyššího hodnocení objektivních metrik dosáhl u této sekvence kodek H.265, subjektivně však byl lépe hodnocen kodek VP9. Jako zajímavost pak působí fakt, že nejvyšší skóre MOS (Obrázek 5-9) sekvence Tall Buildings dosáhla na bitové rychlosti 12 Mbit/s. Tato bitová rychlost se u této sekvence jeví jako nejvýhodnější, dosahuje totiž vysokého objektivního i subjektivního hodnocení při zachování nízké bitové rychlosti. Rozdíly v hodnocení sekvence s datovým tokem 12 a 40 Mbit/s jsou zanedbatelné a dle výsledků MOS pro diváka nepostřehnutelné. Odchylky tvoří řádově desetiny až setiny bodu.



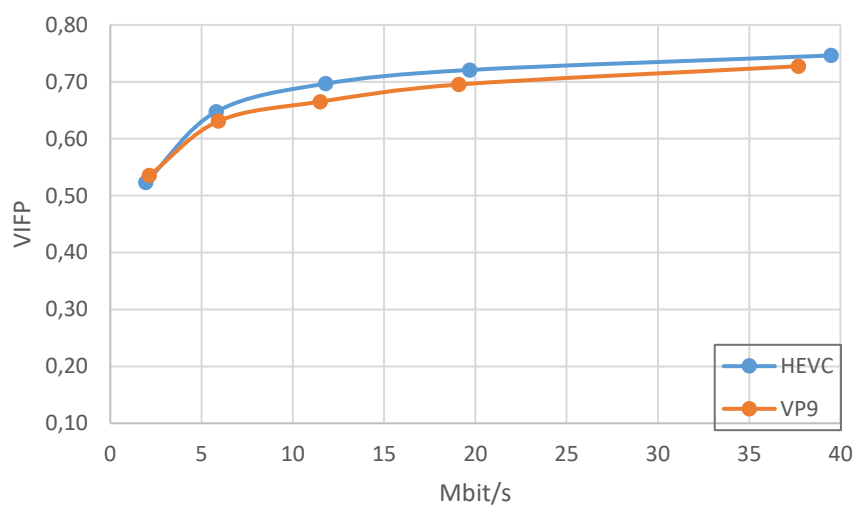
Obrázek 5-5 Tall Buildings UHD průměrné hodnoty PSNR v závislosti na bitové rychlosti



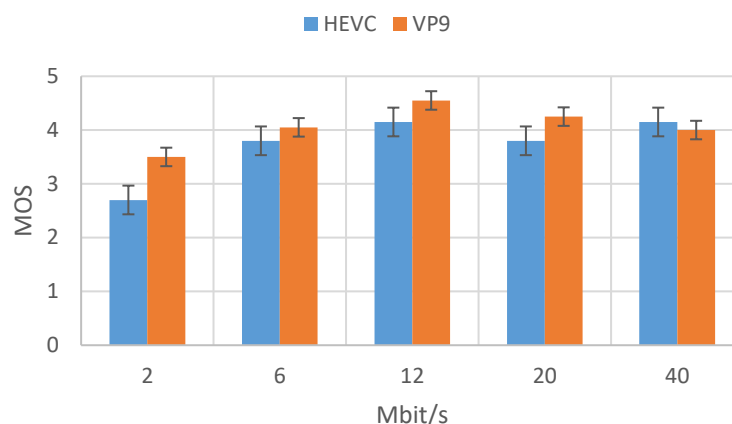
Obrázek 5-6 Tall Buildings UHD hodnoty SSIM v závislosti na bitové rychlosti



Obrázek 5-7 Tall Buildings UHD hodnoty VQM v závislosti na bitové rychlosti



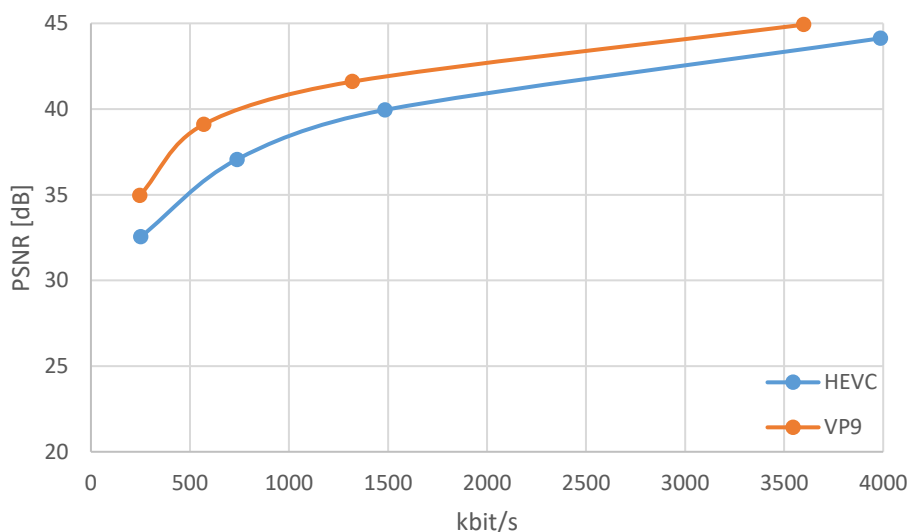
Obrázek 5-8 Tall Buildings UHD hodnoty VIFP v závislosti na bitové rychlosti



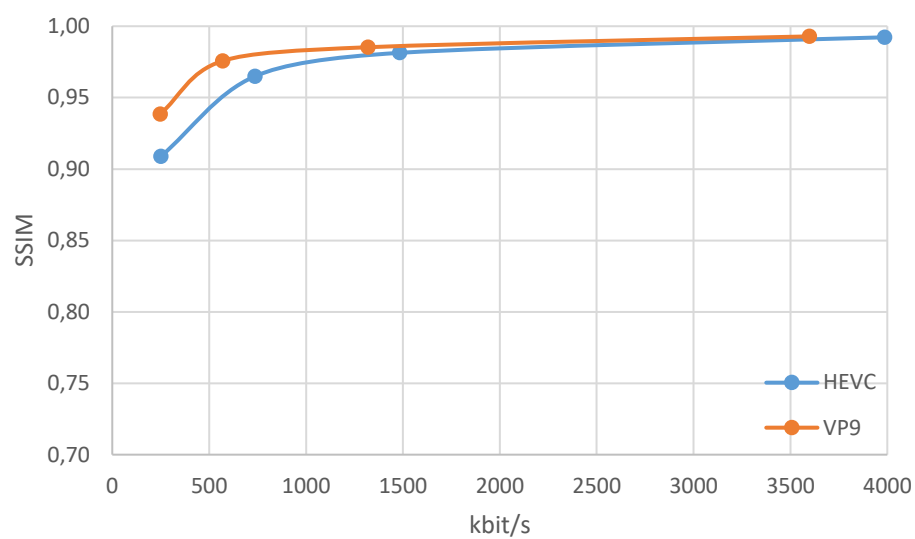
Obrázek 5-9 Tall Buildings UHD průměrné hodnoty MOS

### Nízké rozlišení (856 x 480)

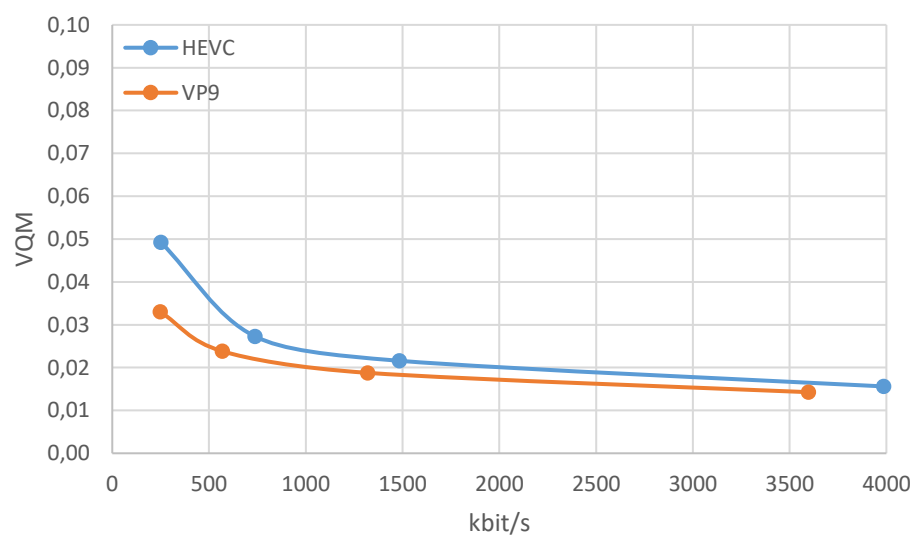
Vyšších hodnot výsledků objektivních metrik hodnocení sekvencí zkomprimovaných v nízkém rozlišení na nízkých bitových rychlostech dosáhl kodek VP9. Dle dosažených výsledků dosahuje kodek VP9 vyšší obrazové kvality spolu s nižší průměrnou bitovou rychlostí. V případě metriky PSNR dosahuje VP9 vysoké hodnoty 41 dB při bitové rychlosti 1300 kbit/s. Za nejvýhodnější z hlediska úspory bitového toku při zachování obrazové kvality lze v tomto případě považovat bitový tok 1000 kbit/s. Při dalším zvyšování bitového toku již není, dle dosažených výsledků, nárůst kvality obdobně strmý. Hodnoty indexu MOS jsou pro oba kodeky téměř identické. Na nižší bitové rychlosti 250 a 750 kbit/s dosáhl vyššího hodnocení kodek VP9, pro 1500 a 2000 kbit/s obdržel vyšší hodnocení kodek HEVC. Rozdíly mezi jednotlivými standardy jsou však řádově desítky bodu a při započítání konfidenčního intervalu lze považovat rozdíly za téměř nepostřehnutelné.



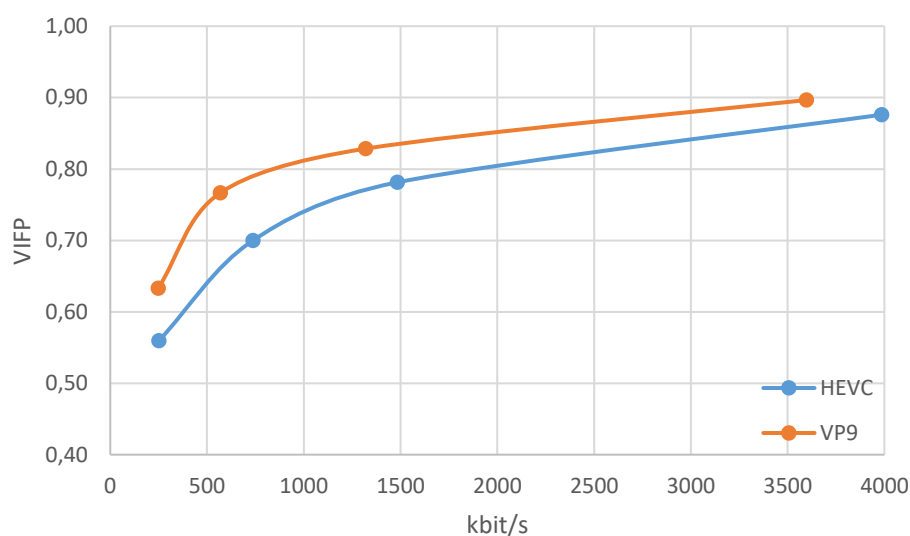
Obrázek 5-10 Tall Buildings SD průměrné hodnoty PSNR v závislosti na bitové rychlosti



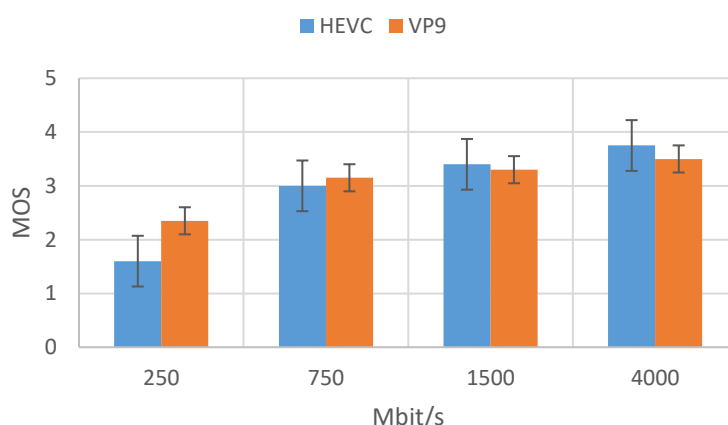
Obrázek 5-11 Tall Buildings SD hodnoty SSIM v závislosti na bitové rychlosti



Obrázek 5-12 Tall Buildings SD hodnoty VQM v závislosti na bitové rychlosti



Obrázek 5-13 Tall Buildings SD hodnoty VIFP v závislosti na bitové rychlosti



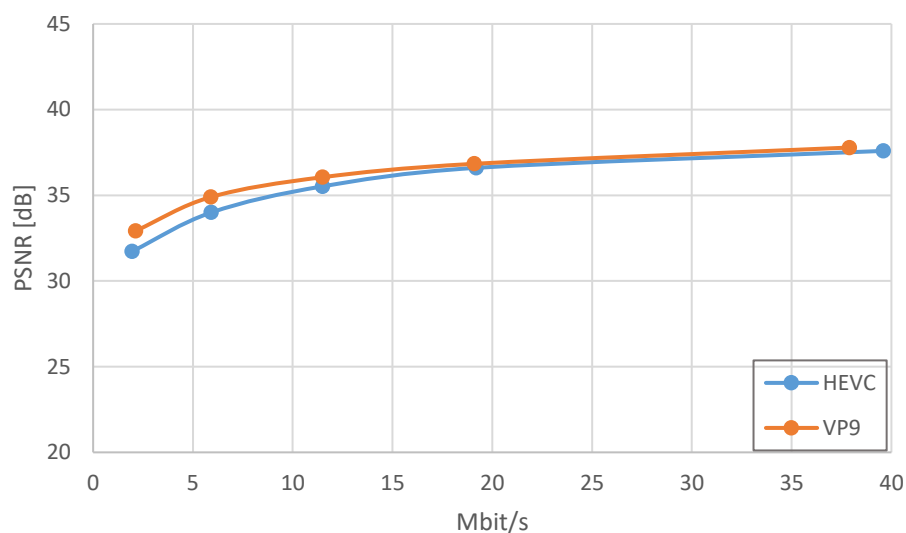
Obrázek 5-14 Tall Buildings SD průměrné hodnoty MOS

## 5.5.2 Sekvence Campfire

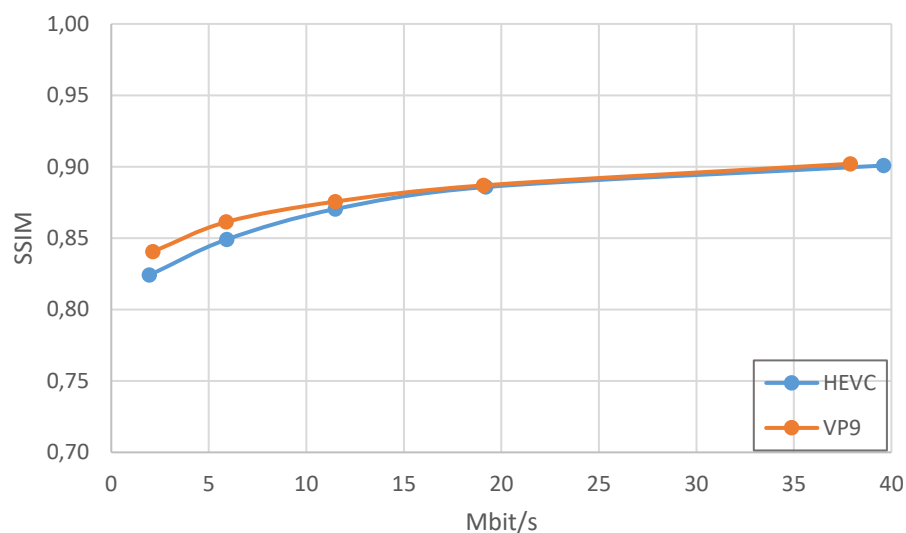
### Rozlišení UHD (3840 x 2160)

Campfire (Tabulka 5-1b) je noční dynamická sekvence s vysokou hodnotou prostorové aktivity TI (Obrázek 5-2). Lze tedy předpokládat, že u takového druhu scény bude k dosažení vysoké obrazové kvality nutný výrazně vyšší bitový tok, než tomu bylo v předchozím případě. Výsledné hodnoty PSNR (Obrázek 5-15) jsou u této sekvence v rozmezí 32 – 37 dB. Z výsledků je patrné, že ani při nejvyšší bitové rychlosti 40 Mbit/s není dle výsledků hodnocení dosaženo stejné obrazové kvality jako v předchozí sekenci. Hodnocení výsledků dalších metrik reflektuje obdobný trend jako u PSNR, nejvyššího hodnocení dosahuje sekvence až při bitové rychlosti 40 Mbit/s. Z hlediska vzájemného srovnání zde mírně převládá kodek VP9, který dosahuje nižší průměrné bitové rychlosti při zachování vysoké obrazové kvality.

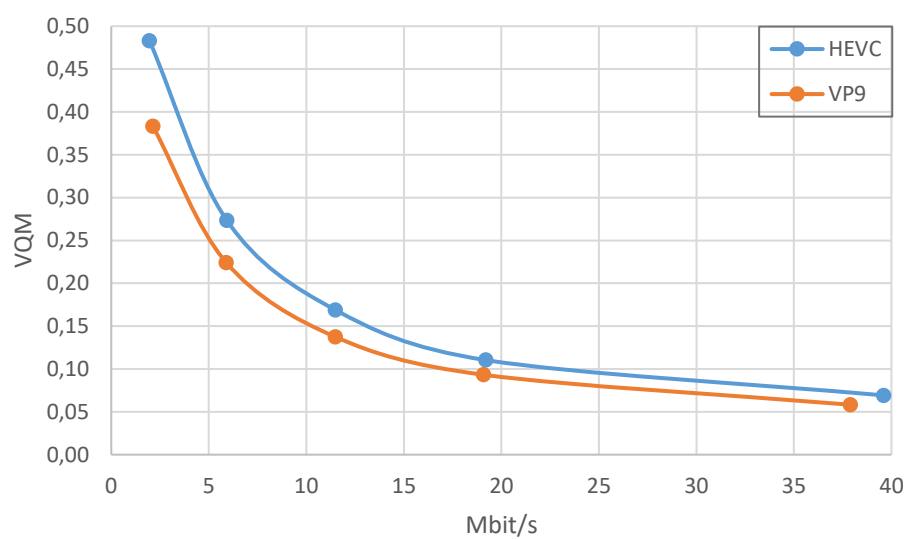
Subjektivní hodnocení pozorovatelů odpovídá objektivnímu měření. Nižší MOS hodnoty jsou získány pro nejnižší bitovou rychlost, kde byl obraz poznamenán viditelnými degradacemi. Udělené skóre pozorovatelů MOS lineárně stoupá spolu s bitovou rychlostí. Důležité je v tomto bodě zmínit, že lepšího hodnocení v subjektivním testu dosáhl kodek H.265. Tato odlišnost mohla být způsobena jednostranným zaměřením pozorovatelů na určitou část scény, v tomto případě oheň a šlehající plameny, na úkor postav v pozadí.



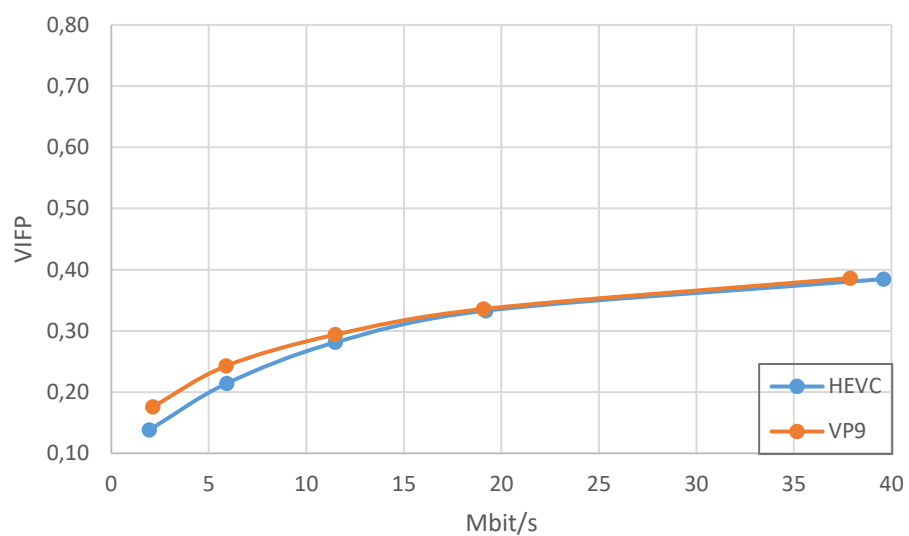
Obrázek 5-15 Campfire UHD průměrné hodnoty PSNR v závislosti na bitové rychlosti



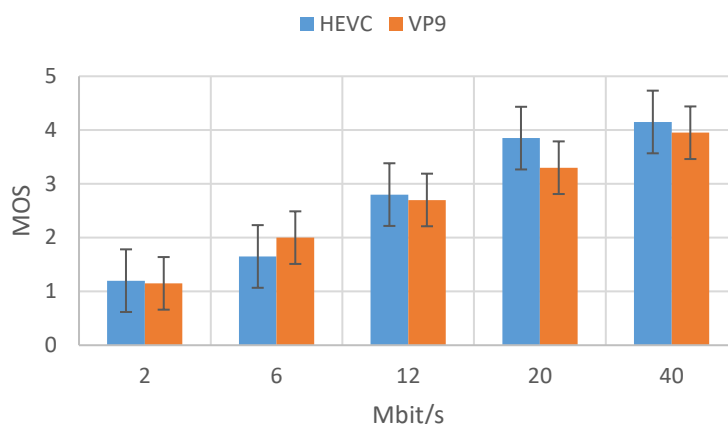
Obrázek 5-16 Campfire UHD hodnoty SSIM v závislosti na bitové rychlosti



Obrázek 5-17 Campfire UHD hodnoty VQM v závislosti na bitové rychlosti



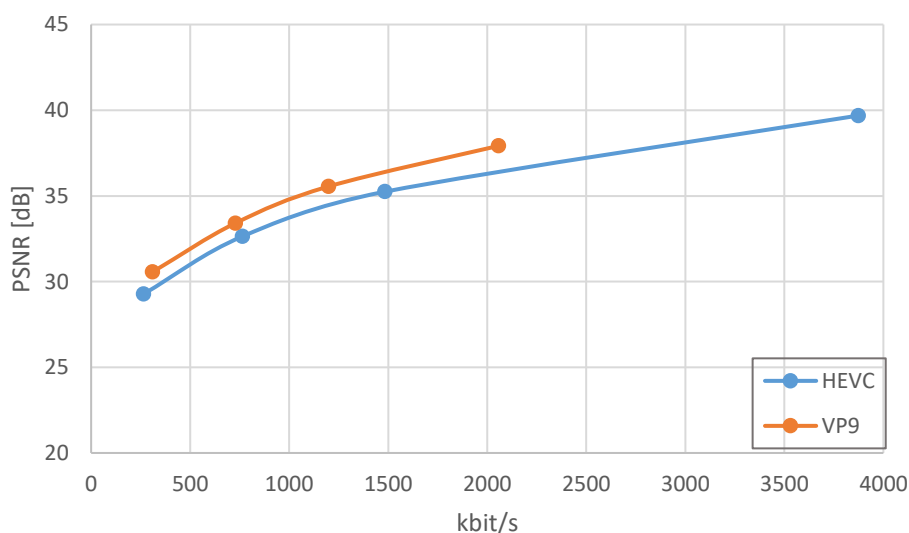
Obrázek 5-18 Campfire UHD hodnoty VIFP v závislosti na bitové rychlosti



Obrázek 5-19 Campfire UHD průměrné hodnoty MOS

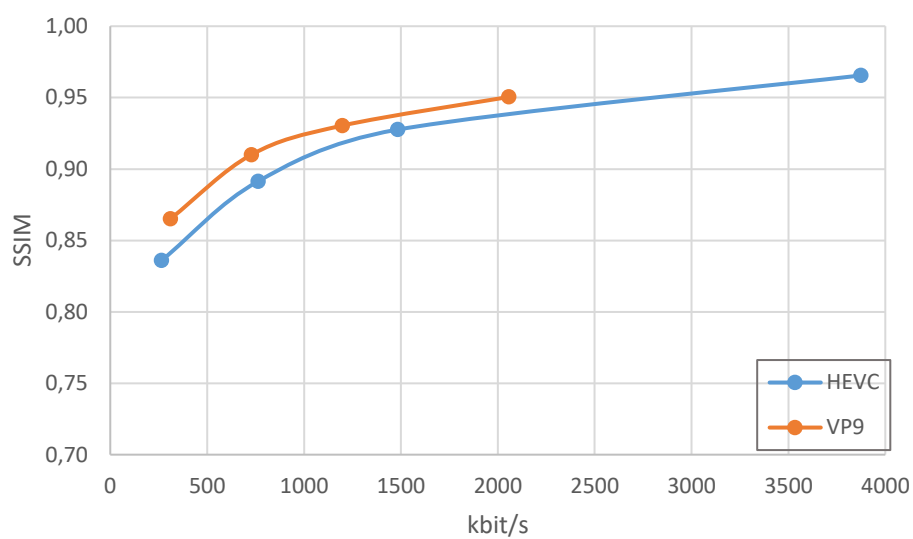
### Nízké rozlišení (856 x 480)

Výsledné hodnoty sekvence Campfire, komprimované v nízkém rozlišení reflektují výsledky hodnocení této scény v UHD rozlišení. Nejvyšších hodnot objektivních metrik dosáhly v tomto případě sekvence s nejvyšší bitovou rychlostí 4000 kbit/s. Rozptyl hodnot PSNR (Obrázek 5-20) byl v tomto případě 29 – 39 dB. Vzhledem k vysokým hodnotám SSIM, VQM i VIFP se jeví jako přijatelná bitová rychlost 1500 kbit/s. Nad tuto hodnotu již nedochází k tak razantnímu nárůstu kvality. Jak je však patrné z výsledných grafů, kodek VP9 při daném nastavení nedokázal docílit maximální bitové rychlosti oproti HEVC. Výsledné sekvence jsou tak zkomprimovány s průměrným bitovým tokem 2056 kbit/s. Na nižších bitových tocích však dosáhl vyššího hodnocení objektivních metrik ve srovnání s HEVC. Prakticky totožný výsledek platí i pro hodnotu MOS.

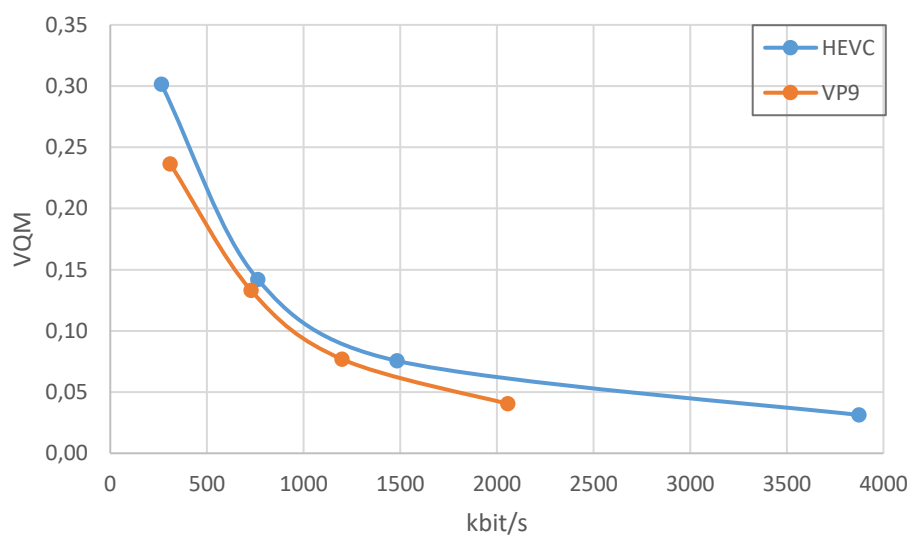


Obrázek 5-20 Campfire SD průměrné hodnoty PSNR v závislosti na bitové rychlosti

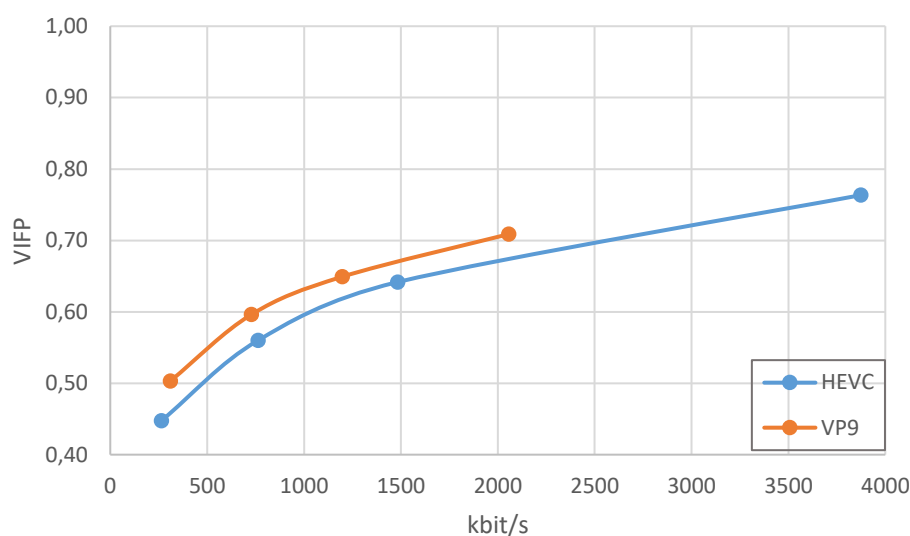




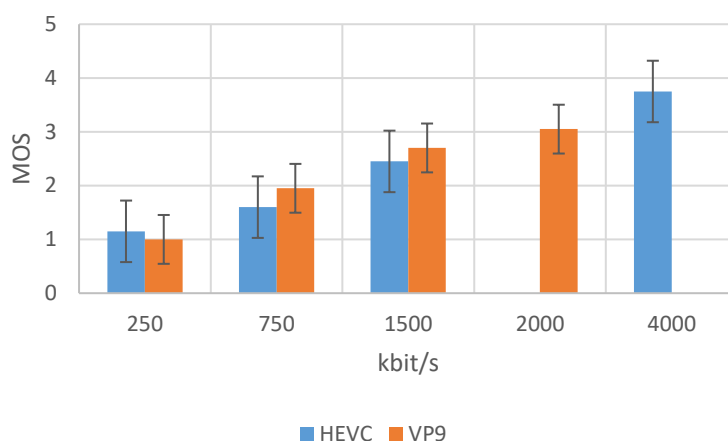
Obrázek 5-21 Campfire SD hodnoty SSIM v závislosti na bitové rychlosti



Obrázek 5-22 Campfire SD hodnoty VQM v závislosti na bitové rychlosti



Obrázek 5-23 Campfire SD hodnoty VIFP v závislosti na bitové rychlosti



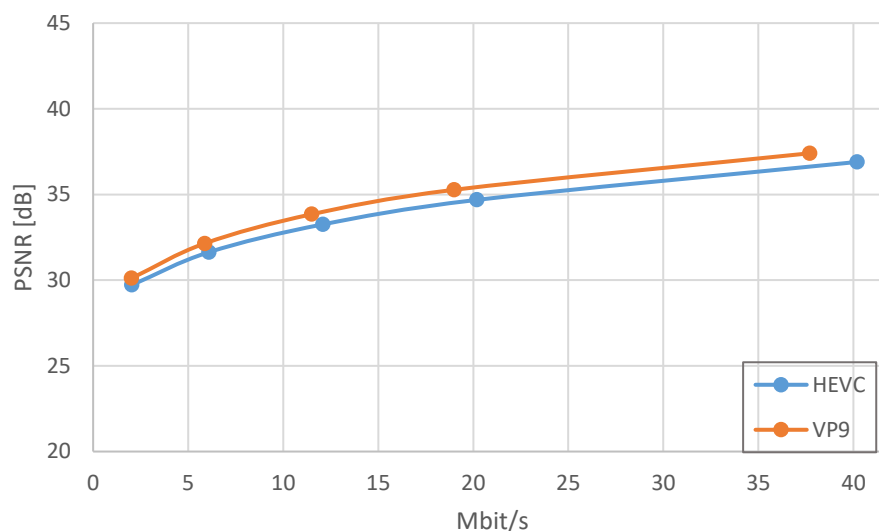
Obrázek 5-24 Campfire SD hodnoty MOS

### 5.5.3 Sekvence Fountains

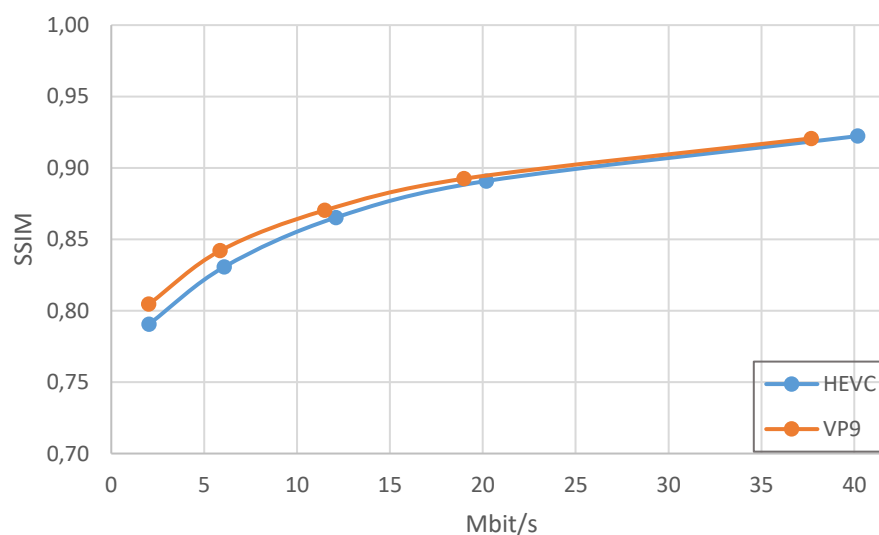
#### Rozlišení UHD (3840 x 2160)

Denní dynamická scéna Fountains (Tabulka 5-1c) která i přes své poměrně nízké hodnoty časové i prostorové aktivity (Obrázek 5-2) představuje druh scény požadující vysokou hodnotu bitového toku. Nejnáročnější část obrazu pro zpracování v tomto případě byla stříkající voda z fontány a také vodní hladina čerící se pod dopadem drobných kapek. V této části obrazu bylo možné pozorovat markantní degradace při nízké bitové rychlosti. Výsledné hodnoty PSNR (Obrázek 5-25) se pohybují v rozmezí 29 – 37 dB. Nejvyššího hodnocení SSIM, VQM i VIFP v tomto případě dosáhly sekvence s nejvyšším bitovým tokem 40 Mbit/s. Při vzájemném srovnání obou standardů dosahuje na nižších bitových rychlostech (2 a 6 Mbit/s) lepších výsledků kodek VP9.

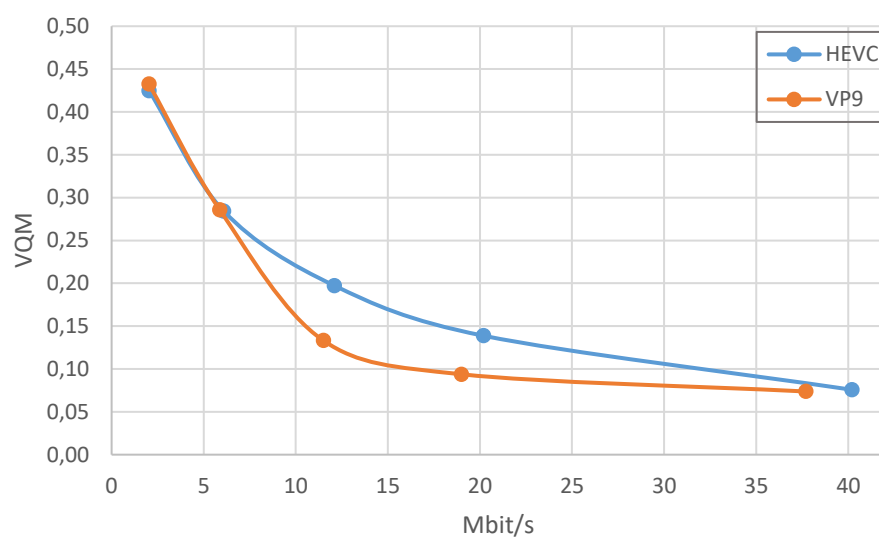
Zbývající výsledky jsou téměř identické pro oba standardy. Zajímavostí tedy je, že i v tomto případě obdržel kodek H.265 vyšší skóre MOS (Obrázek 5-29) při subjektivním hodnocení pozorovatelů. Také v tomto případě může být výsledek ovlivněný koncentrací diváka na dynamickou část scény, jejíž zpracování dle dostupných srovnání [3] lépe zvládá HEVC. Vyšší skóre MOS obdržel standard VP9 pouze u sekvence o nejnižší bitové rychlosti 2 Mbit/s.



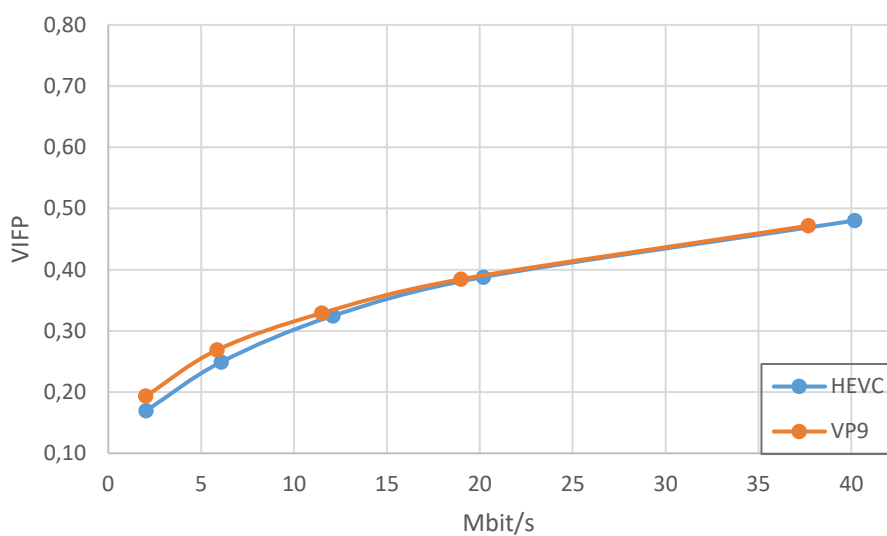
Obrázek 5-25 Fountains UHD Průměrné hodnoty PSNR v závislosti na bitové rychlosti



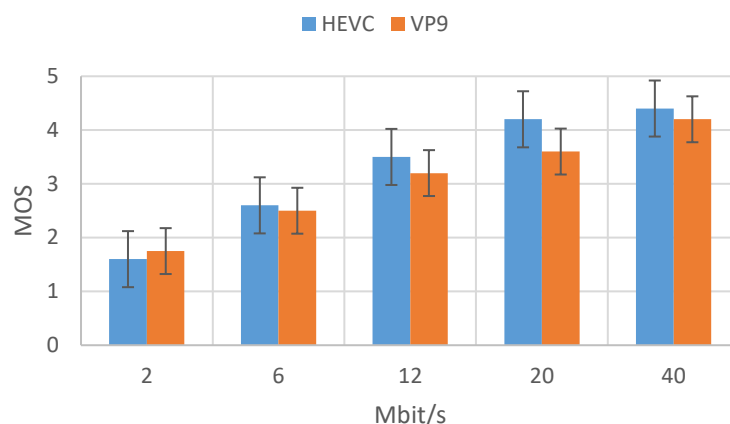
Obrázek 5-26 Fountains UHD hodnoty SSIM v závislosti na bitové rychlosti



Obrázek 5-27 Průměrné hodnoty VQM v závislosti na bitové rychlosti



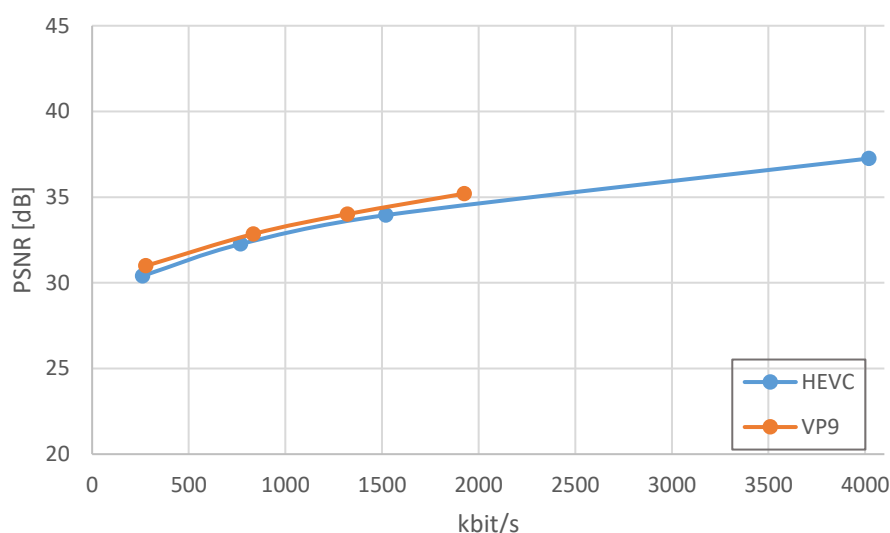
Obrázek 5-28 Fountains UHD hodnoty VIFP v závislosti na bitové rychlosti



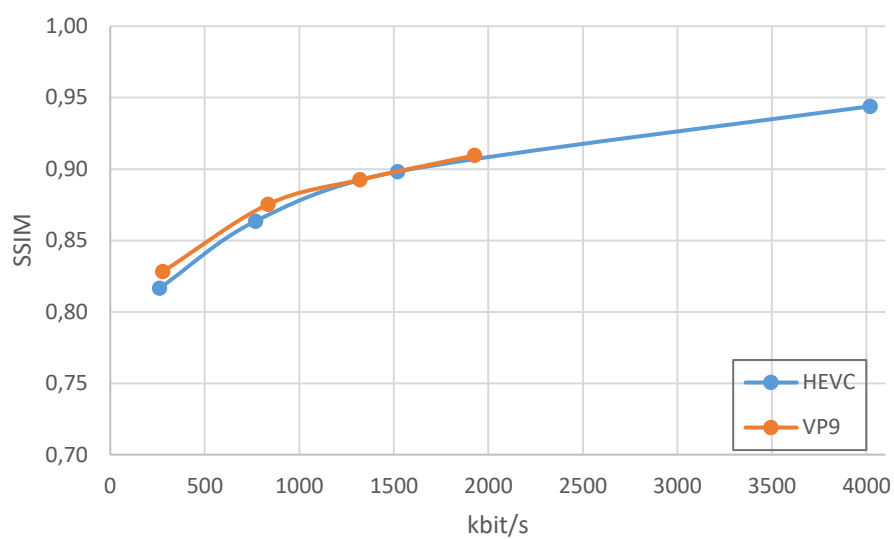
Obrázek 5-29 Fountains UHD průměrné hodnoty MOS

### Nízké rozlišení (856 x 480)

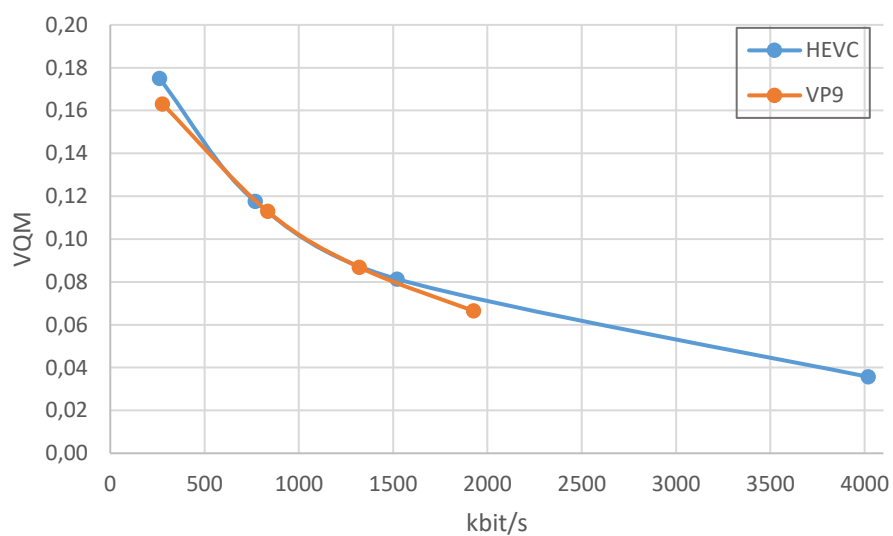
Výsledky sekvence Fountains komprimované v nízkém rozlišení vykazují podobné vlastnosti jako při UHD rozlišení. Průměrné hodnoty PSNR (Obrázek 5-30) se pohybují v rozmezí 27 – 38 dB. Rozdíly ve výsledných hodnotách jednotlivých metrik jsou mezi VP9 a HEVC zanedbatelné. Nejvyššího hodnocení dosahuje kodek HEVC při bitové rychlosti 4000 kbit/s. Podobně jako v předchozím případě, kodek VP9 nedosáhl nejvyšší požadované bitové rychlosti. Zjištěná bitová rychlost je 1926 kbit/s oproti 4020 kbit/s u HEVC. Při subjektivním hodnocení kvality této scény pozorovatelé korektně identifikovali vzrůstající obrazovou kvalitu s rostoucí bitovou rychlostí. Vyšší skóre MOS (Obrázek 5-34) však i v tomto případě obdržel kodek HEVC. Průměrné hodnocení každé sekvence bylo o 0,5 bodu vyšší oproti VP9.



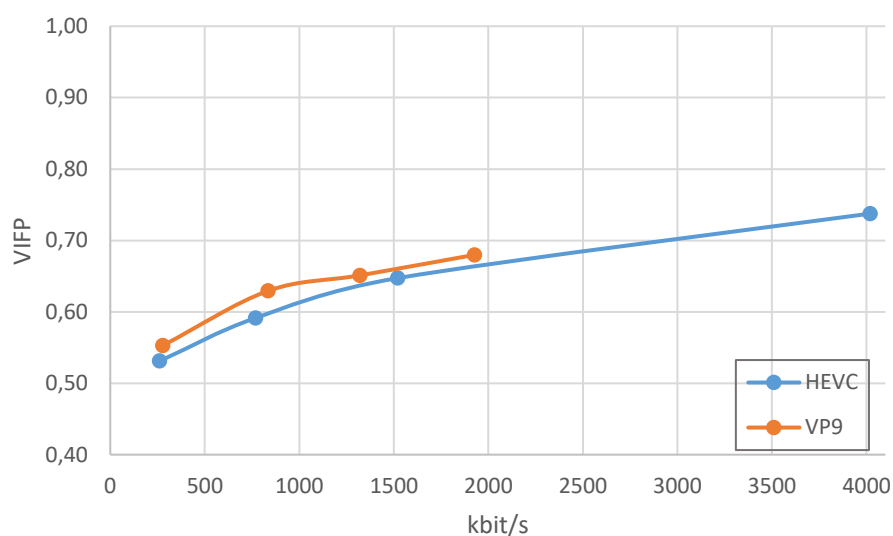
Obrázek 5-30 Fountains SD průměrné hodnoty PSNR v závislosti na bitové rychlosti



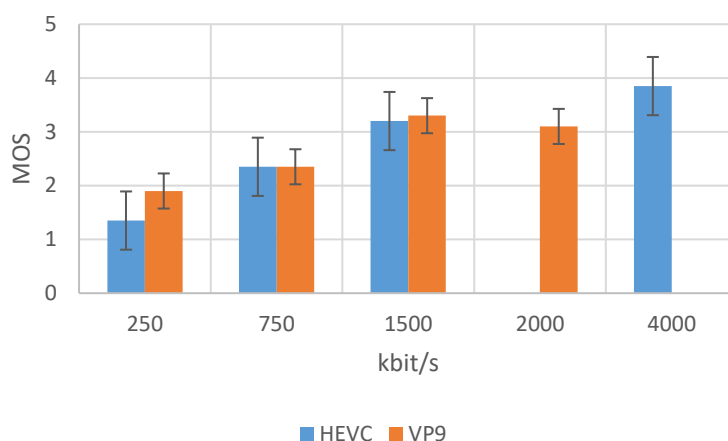
Obrázek 5-31 Fountains SD hodnoty SSIM v závislosti na bitové rychlosti



Obrázek 5-32 Fountains SD hodnoty VQM v závislosti na bitové rychlosti



Obrázek 5-33 Fountains SD hodnoty VIFP v závislosti na bitové rychlosti



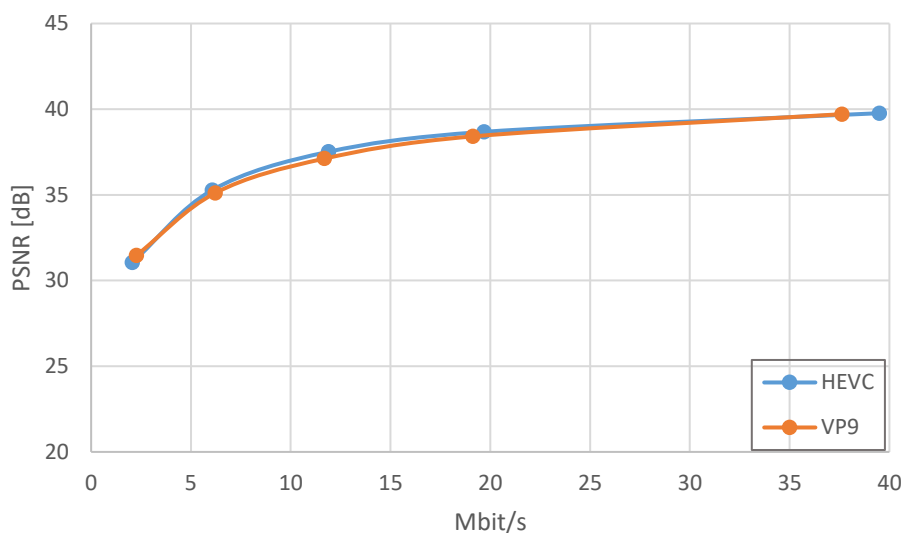
Obrázek 5-34 Fountains SD průměrné hodnoty MOS

## 5.5.4 Sekvence Woods

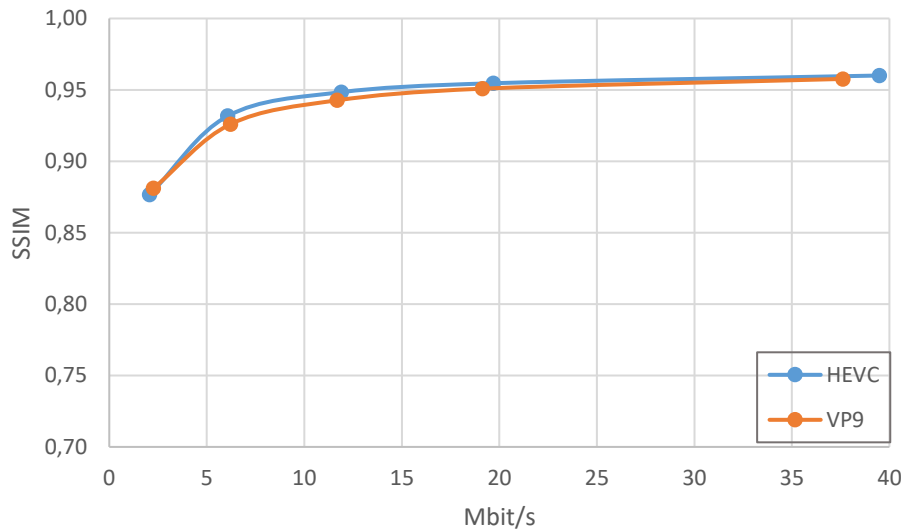
### Rozlišení UHD (3840 x 2160)

Videosekvence Woods (Tabulka 5-1d) obsahuje denní scénu s vysokými hodnotami časové i prostorové aktivity. Jde o komplexní scénu s plynulým pohybem kamery. Podobně jako Buildings (Tabulka 5-1a) je i tato scéna vhodná pro plné využití predikčních vlastností jednotlivých kodeků. Výsledné hodnoty metriky PSNR (Obrázek 5-35) se pohybují v rozmezí 31 – 39 dB. Z hlediska výsledků metrik SSIM, VIFP i VQM lze za nejvhodnější považovat bitovou rychlost 20 Mbit/s, neboť při tomto bitovém toku dosahuje vysoké obrazové kvality (hodnota VQM 0,028 pro HEVC i VP9). Zvýšení bitového toku na dvojnásobek, tj. 40 Mbit/s v tomto případě přináší zanedbatelné zvýšení obrazové kvality (hodnota VQM 0,025 pro HEVC i VP9). Z hlediska vzájemného srovnání kodeků jsou výsledky VP9 i HEVC prakticky totožné.

Tento závěr koresponduje i s výsledky subjektivního hodnocení, nejvyšší skóre MOS (Obrázek 5-39) obdržel kodek VP9 pro bitovou rychlost 12 Mbit/s. Udělené hodnocení bylo v průměru o 0,5 bodu vyšší. Zbylé výsledky MOS jsou obdobně jako objektivní skóre, prakticky totožné.

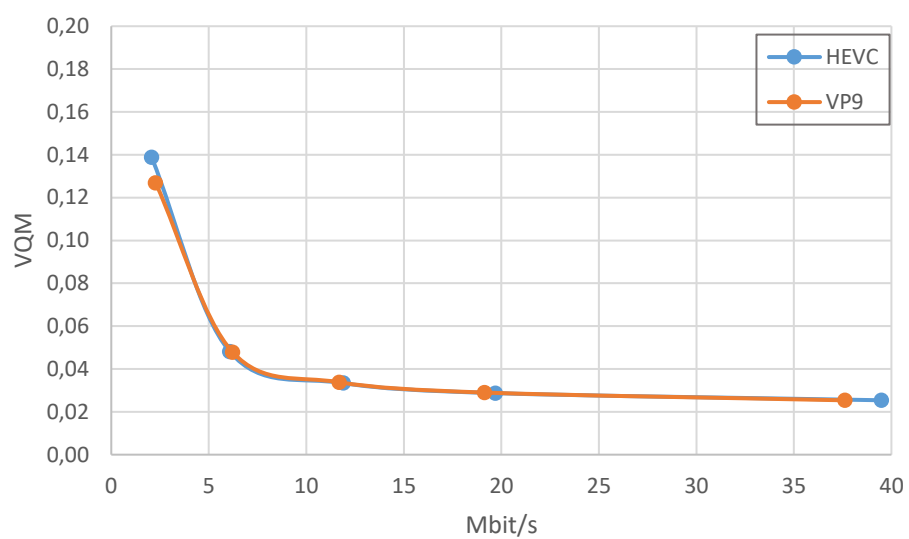


Obrázek 5-35 Woods UHD průměrné hodnoty PSNR v závislosti na bitové rychlosti

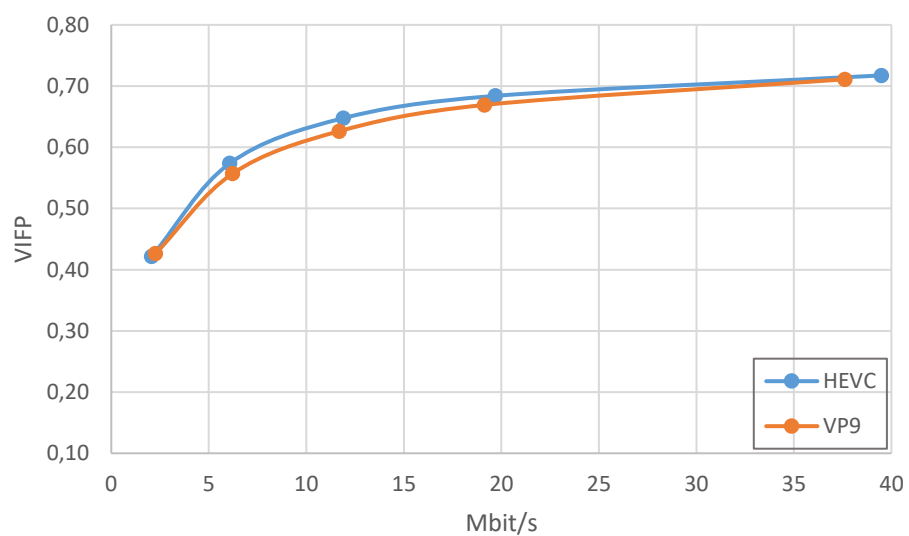


Obrázek 5-36 Woods UHD hodnoty SSIM v závislosti na bitové rychlosti

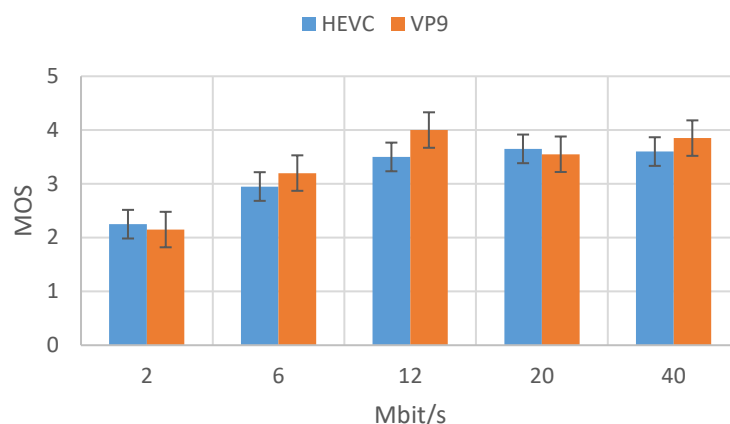




Obrázek 5-37 Woods UHD hodnoty VQM v závislosti na bitové rychlosti



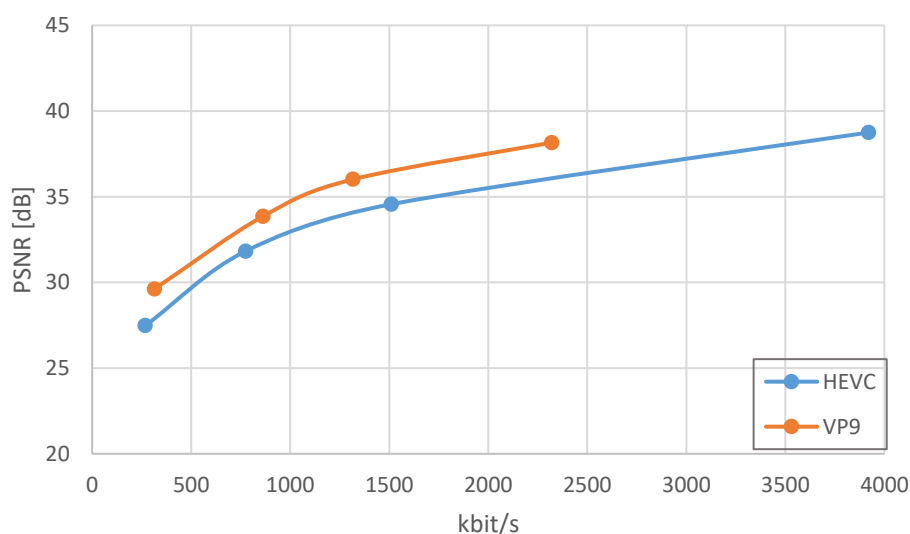
Obrázek 5-38 Woods UHD hodnoty VIFP v závislosti na bitové rychlosti



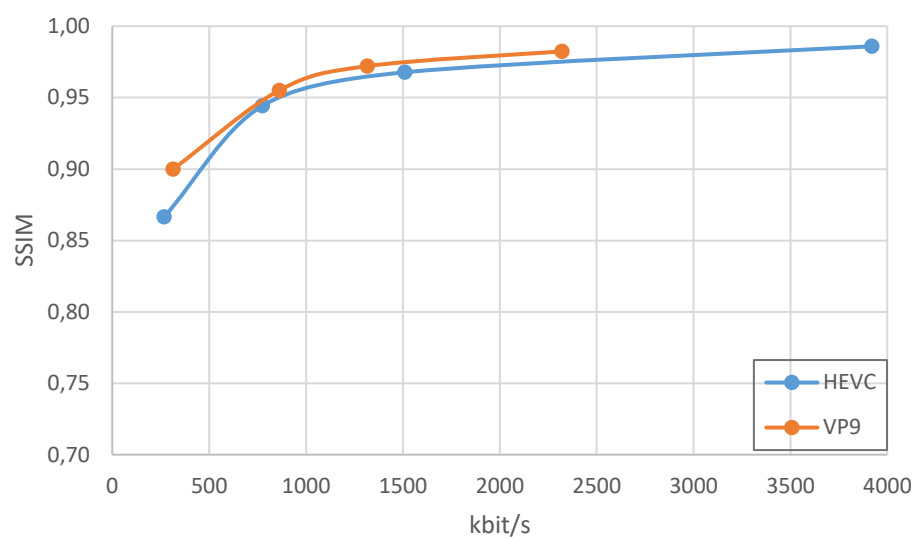
Obrázek 5-39 Woods UHD průměrné hodnoty MOS

### Nízké rozlišení (856 x 480)

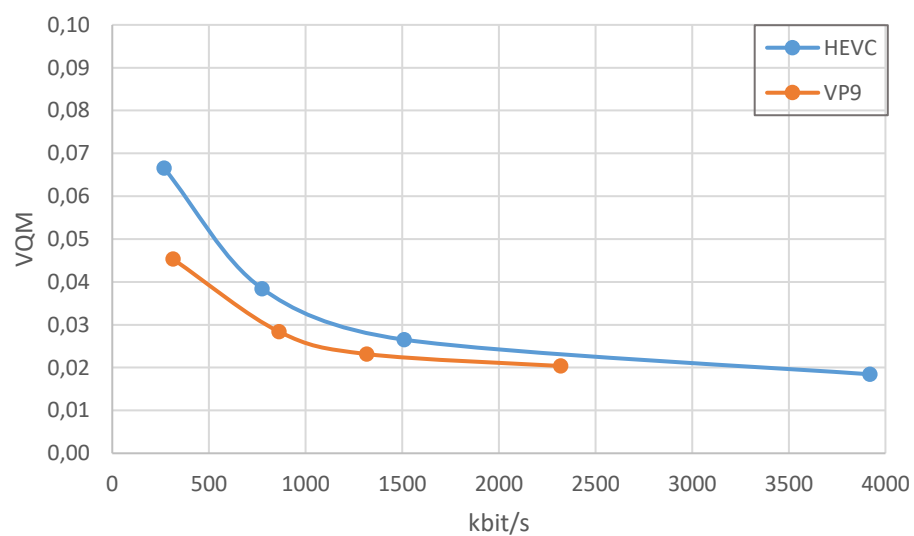
Výsledky sekvece Woods v nízkém rozlišení kopírují trend vyššího hodnocení kodeku VP9 pro tento typ videa. Důležité je v tomto případě zmínit fakt, že i přes nedodržení nejvyšší požadované bitové rychlosti (4000 kbit/s) v případě kodéru VP9, který sekvence zkomprimoval s bitovým tokem 2320 kbit/s, není tento nedostatek při zkoumání výše uvedenými objektivními metodami téměř patrný, neboť oba kodéry dosahují téměř totožných výsledných hodnot (hodnoty VQM 0,018 HEVC a 0,02 VP9). S tímto tvrzením však nekorespondují výsledky MOS (Obrázek 5-44). Výrazně vyššího hodnocení dosáhl při nejvyšším bitovém toku kodek HEVC s průměrnou známkou 3,85 ve srovnání s 3,1 pro VP9. Z dosavadních výsledků je tedy možné vyvodit závěr, že u sekvencí s nízkým rozlišením pozorovatel lépe rozezná rozdíly mezi různými bitovými toky ve srovnání s UHD rozlišením na velké obrazovce, kde jsou jednotlivé rozdíly hůře postřehnutelné.



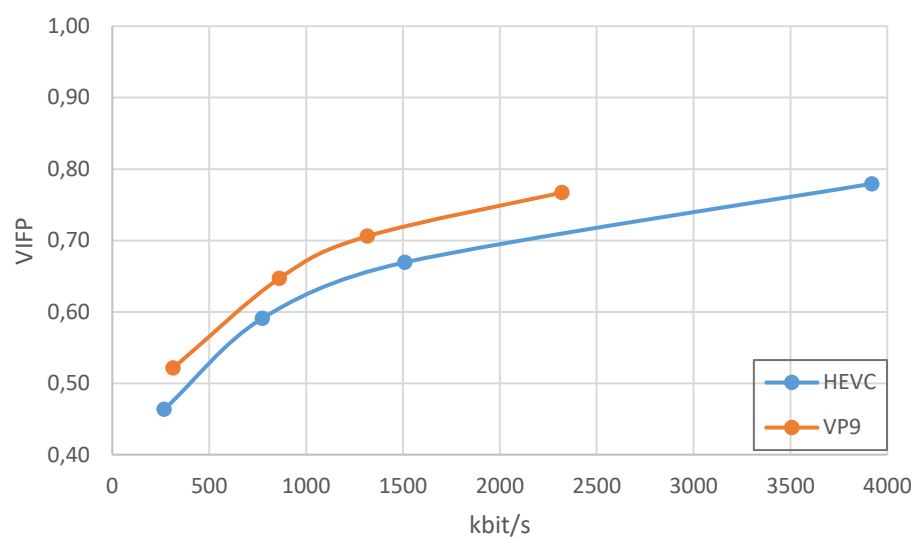
Obrázek 5-40 Woods SD průměrné hodnoty PSNR v závislosti na bitové rychlosti



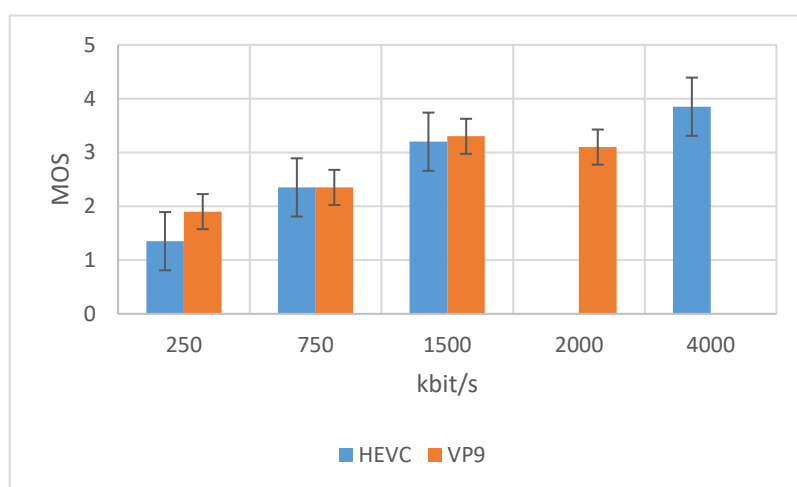
Obrázek 5-41 Woods SD hodnoty SSIM v závislosti na bitové rychlosti



Obrázek 5-42 Woods SD hodnoty VQM v závislosti na bitové rychlosti



Obrázek 5-43 Woods SD Průměrné hodnoty VIFP v závislosti na bitové rychlosti



Obrázek 5-44 Woods SD průměrné hodnoty MOS

## 5.6 Korelační analýza výsledků

Pro zjištění míry statistické závislosti mezi objektivními metrikami a subjektivním hodnocením je vhodné provést výpočet korelačních koeficientů jednotlivých metrik. Za tímto účelem byl využit Spearmanův korelační koeficient (Spearman's rank order correlation coefficient) SROCC, který se počítá dle rovnice [27]:

$$SROCC = 1 - \frac{6 \sum_i (p_i - q_i)^2}{n(n^2 - 1)}, \quad (5.4)$$

kde hodnoty  $p_i$  a  $q_i$  reprezentují vzestupnou hodnotu pořadí jednotlivých výsledků,  $n$  reprezentuje celkový počet vzájemně srovnávaných množin hodnot. Výsledné hodnoty korelačního koeficientu se pohybují od -1 do 1. Nulové hodnoty dosahuje koeficient dvou množin, mezi kterými není žádný vztah. To znamená, že vzájemně nekorelují. Naopak maximální hodnoty (-1 nebo 1) dosáhne koeficient pro vzájemně monotónní množiny.

Korelační koeficienty byly počítány mezi hodnotami MOS a jednotlivými výsledky objektivních metrik, zvláště pro nízké a vysoké rozlišení. Výpočet byl proveden pomocí programu MATLAB. Výsledné korelační koeficienty pro jednotlivé metriky jsou uvedeny v Tabulce 5-2. Z výsledků je patrné, že vyšší míry korelace bylo dosaženo u sekvencí s nízkým rozlišením. V případě metriky VQM byla použita absolutní hodnota výsledného koeficientu, neboť vysoká hodnota MOS značí vysokou kvalitu sekvence, kdežto hodnoty VQM se stoupající kvalitou klesají. Výsledné korelační koeficienty by tedy byly záporné. Nejvyšší míry korelace dosáhly dle dostupných studií [25] metriky VIFP a VQM, které jsou nejvhodnější pro vyhodnocení kvality kódovaných sekvencí.

Tabulka 5-2 Korelační koeficienty objektivních metrik

		Objektivní metriky			
Korelace	Rozlišení	PSNR	SSIM	VIFP	VQM
SROCC	UHD	0,77	0,79	0,72	0,83
	SD	0,86	0,87	0,94	0,79

# ZÁVĚR

Cílem této práce bylo prostudovat a formulovat vlastnosti moderních kódovacích standardů VP9 a H.265. V první kapitole byly popsány vlastnosti a funkce jednotlivých kódovacích standardů. Další část práce byla věnována možnostem hodnocení kvality komprimovaných sekvencí pomocí objektivních a subjektivních metrik. Z objektivních metrik byly pro účely práce vybrány plně a částečně referenční metriky PSNR, SSIM, VQM a VIFP. Subjektivní metriky byly uvažovány ACR a DSCQS. Závěrečná část se zaměřuje na popis postupů komprese videa s využitím dostupných referenčních implementací jednotlivých kodeků.

Pro experimentální část práce byly vybrány 4 krátké videosekvence v UHD rozlišení, které byly následně komprimovány oběma kodeky s nastavením různých bitových rychlostí v rozmezí 2 – 40 Mbit/s. Tyto 4 sekvence byly následně bezztrátově převedeny do nízkého rozlišení a komprimovány oběma kodeky s nastavením bitového toku v rozsahu 250 – 4000 kbit/s. Celkem tak bylo získáno 40 sekvencí v UHD rozlišení a 32 sekvencí v nízkém rozlišení (SD). Na tyto sekvence byly následně aplikovány výše popsané objektivní metriky.

Dále byly uspořádány subjektivní testy hodnocení kvality pro UHD a SD sekvence. Byla zvolena metoda ACR, jejíž provedení bylo realizováno dle standardizovaných postupů. Pro toto hodnocení byly připraveny dvě sady sekvencí pro UHD a SD rozlišení. Jednotlivé sady sekvencí byly přehrány na odlišných zobrazovačích (UHD TV a Full HD TV). Získaná data byla statisticky zpracována a následně porovnána s výsledky objektivních metrik. Výsledky jednotlivých hodnocení byly zaneseny do grafů v závislosti na bitové rychlosti. Následně byl vypočten Spearmanův korelační koeficient za účelem zjištění míry korelace jednotlivých objektivních metrik vůči subjektivnímu hodnocení. Ze zjištěných hodnot vyplývá, že vyšší míry korelace dosáhly komplexní metriky VQM a VIFP.

Z dosažených výsledků vyplývá, že oba uvedené standardy dokáží velmi efektivně redukovat bitový tok videa ve vysokém rozlišení. Po vzájemném srovnání jednotlivých parametrů je možné usuzovat, že VP9 dosahuje vyšší úspory bitového toku, zejména u sekvencí v nízkém rozlišení, zatímco HEVC dosahuje vyššího hodnocení při zpracování videa v UHD rozlišení, a to zejména v hodnocení subjektivním.

Pro další rozšíření této práce by bylo vhodné provést podrobnou analýzu nastavení jednotlivých parametrů kódování, zejména pak jejich vliv na výsledný bitový tok. Proměnné nastavení parametrů by následně mohlo být otestováno na rozšířené sadě sekvencí a následně provedeno vyhodnocení pomocí uvažovaných metrik.

# LITERATURA

- [1] MUKHERJEE, D., BANKOSKI, J., GRANGE, A. *The latest open-source video codec VP9 – An overview and preliminary results*. USA: IEEE, 2013. ISBN 978-1-4799-0292-7 [cit. 15.12.2015]
- [2] AKRAMULLAH, S. *Digital video concepts, methods, and metrics: quality, compression, performance, and power trade-off analysis*. 2013, 344 s.
- [3] SHARABAYKO, M. P., PONOMAREV, O. G., CHERNYAK R. I. *Intra Compression Efficiency in VP9 and HEVC*. Russia: Hikari Ltd, 2013. ISSN 6803 – 6824. [cit.15.12.2015] Dostupné z <http://m-hikari.com/ams/ams-2013/ams-137-140-2013/sharabaykoAMS137-140-2013.pdf>
- [4] SULLIVAN, G. J., OHM, T., WOO-JIN, H., WIEGAND, T. *Overview of the High Efficiency Video Coding (HEVC) Standard*. *Circuits and Systems for Video Technology, IEEE Transactions on* [online]. USA: IEEE, 0121n. 1., 22(12): 1649-1668 [cit. 16.12.2015]. ISSN 1051-8215.
- [5] DOMINGUEZ, H., VERGARA VILLEGAS, O., CRUZ SANCHEZ, V. *The H.264 Video Coding Standard*. USA: IEEE, 1403, [cit. 2017-05-22]. ISSN 0278-6648.
- [6] IL-KOO, K., SUNIL, L., PIAO, Y., CHEN, J. *Coding efficiency comparison of new video coding standards: HEVC vs VP9 vs AVS2 video*. USA: IEEE, 2014 [cit. 15.12.2015]. ISSN 1945-7871.
- [7] HANUS, S. *Základy televizní techniky, přednášková prezentace č. 10* Brno: Vysoké učení technické, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky, 2015.
- [8] ITU-T Rec. P.910. *Subjective video quality assessment methods for multimedia applications*. Geneva: ITU, 2008.
- [9] ITU-T Rec. BT.500 - 13. *Methodology for the subjective assessment of the quality of television pictures*. Geneva: ITU, 2012.
- [10] ZHOU, W., BOVIK, A.C., SHEIKH, H.R., SIMONCELLI, E.P. *Image quality assessment: from error visibility to structural similarity*. USA: IEEE, 2004. ISSN 1057-7149. [cit 16.12.2015]
- [11] ITU-T Rec. J.144. *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*. Geneva: ITU, 2004.
- [12] ITU-T Rec. J.244. *Full reference and reduced reference calibration methods for video transmission systems with constant misalignment of spatial and temporal domains with constant gain and offset*. Geneva: ITU, 2008. [cit. 22.5.2017]
- [13] ITU-T Rec. J.249. *Perceptual video quality measurement techniques for digital cable television in the presence of a reduced reference*. Genova: ITU 2010. [cit. 22.5.2017]
- [14] PINSON, M.H. a S. WOLF. *A new standardized method for objectively measuring video quality*. USA: IEEE, 0409n. 1., 50(3), 312-322 [cit. 2017-05-22]. DOI: 10.1109/TBC.2004.834028. ISSN 0018-9316.
- [15] SHEIKH, H. R., BOVIK, A.C. *Image information and visual quality*. USA: IEEE, 2006. ISSN 1057-7149. [cit 16.12.2015]

- [16] The WebM Project | Developer Overview . *The WebM Project | Welcome to the WebM Project* [online]. Dostupné z: <https://www.webmproject.org/code/>
- [17] MinGW | Minimalist GNU for Windows. *MinGW | Minimalist GNU for Windows* [online]. Copyright © 2017 [cit. 23.05.2017]. Dostupné z: <http://www.mingw.org/>
- [18] High Efficiency Video Coding (HEVC) | JCT-VC. *High Efficiency Video Coding (HEVC) / JCT-VC* [online]. Copyright © 2013 [cit. 23.05.2017]. Dostupné z: <https://hevc.hhi.fraunhofer.de/>
- [19] BOSSEN, F., FLYNN, D., SHARMAN, K., SÜHRING, K. *HM Software Manual*. 2016. [cit. 23.05.2017]. Dostupné z <https://hevc.hhi.fraunhofer.de/>
- [20] Builds - Zerano FFmpeg. *301 Moved Permanently* [online]. Copyright © 2017 Kyle Schwarz [cit. 23.05.2017]. Dostupné z: <http://ffmpeg.zerano.com/builds/>
- [21] VQMT: Video Quality Measurement Tool | MMSPG. *MMSPG / EPFL* [online]. Copyright © [cit. 23.05.2017]. Dostupné z: <http://mmspg.epfl.ch/vqmt>
- [22] Description of VQM Tools - ITS. *Institute for Telecommunication Sciences - ITS* [online]. Dostupné z: <https://www.its.bldrdoc.gov/resources/video-quality-research/guides-and-tutorials/description-of-vqm-tools.aspx>
- [23] SONG, L., TANG, X., ZHANG, W., YANG, X., XIA, P. *The SJTU 4K Video Sequence Dataset*, the Fifth International Workshop on Quality of Multimedia Experience (QoMEX2013), Klagenfurt, Austria, 2013.
- [24] The WebM Project | VP8 Encode Parameter Guide . *The WebM Project | Welcome to the WebM Project* [online]. Dostupné z: <http://www.webmproject.org/docs/encoder-parameters/#2-pass-best-quality-vbr-encoding>
- [25] KUFA, J.; POLÁK, L.; KRATOCHVÍL, T. *HEVC/H.265 vs. VP9 for Full HD and UHD Video: Is There Any Difference in QoE?* In *Proceedings ELMAR- 2016*. Zagreb: Faculty of Electrical Engineering and Computing, 2016. s. 51-55. ISBN: 978-953-184-221- 1.
- [26] Ishihara Color Test – Color Blindness. *Color Blindness – Information : Identification : Solutions :: Everything you need to know about Colour Blindness* [online]. Dostupné z: <http://www.colour-blindness.com/colour-blindness-tests/ishihara-colour-test-plates/>
- [27] HANHART, P., KORSHUNOV, P., EBRAHIMI, T. *Benchmarking of quality metrics on ultra-high definition video sequences*. In: *Digital Signal Processing (DSP)*, 2013 18th International Conference on [online]. IEEE, 1307, [cit. 21.5.2017]. DOI: 10.1109/ICDSP.2013.6622760. ISSN 1546-1874.
- [28] FENDRICH, V. *Comparision of VP9 and H.265 Encoding Standards*. In *Proceedings of the 23nd Conference STUDENT EEICT 2017*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2017. s. 112 – 114. ISBN 978-80-214-5496-5.



# SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

ACR	Absolute Category Rating
ADST	Asymetric Discrete Sine Transform
AVC	Advanced Video Coding
CABAC	Context Adaptive Binary Arithmetic Coding
CB	Coding Block
CSV	Comma Separated Value
CTU	Code Tree Unit
DCS	Differential Category Rating
DCT	Discrete Cosine Transform, diskrétní kosinová transformace.
DFT	Discrete Fourier Transform, diskrétní Fourierova transformace.
DPB	Decoded Picture Buffer
DSCQS	Double Stimulus Continous Quality Scale
DVQ	Digital Video Quality
DWT	Discrete Wavelet Transform
HD	High Definition
HEVC	High Efficiency Video Coding
HVC	Human Visual System Model
ITU	International Telecommunication Union
MOS	Mean Opinion Score
MPEG	Moving Picture Experts Group
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural Similarty Index
UHD	Ultra High Definition
VCEG	Video Coding Experts Group
VQM	Video Quality Model
WHT	Walsh-Hamard Transform

# A LABORATORNÍ ÚLOHA

## A.1 Návrh laboratorní úlohy

Digitální televizní a rozhlasové systémy (MDTV, LDTV)

Laboratorní úloha

### **SROVNÁNÍ STANDARDŮ VP9 A H.265 PRO KÓDOVÁNÍ VIDEOSEKVENCÍ**

Účelem této úlohy je vzájemné srovnání dvou nejnovějších kompresních standardů VP9 a H.265 pro kódování videosekvencí, zejména ve vysokém rozlišení. Toto srovnání bude provedeno prostřednictvím několika krátkých videí, zkomprimovaných oběma standardy s různým nastavením bitového toku. Kvalita takto zkomprimovaných sekvencí bude vyhodnocena prostřednictvím dvou standardizovaných objektivních metrik PSNR a VQM. Následně si vyzkoušíte také subjektivně zhodnotit kvalitu těchto sekvencí a vzájemně porovnáte výsledky objektivních metrik a subjektivního hodnocení kvality.

#### **Teoretický úvod**

Kodek s označením VP9 představuje novou generaci video kodeku, vyvíjeného v rámci projektu WebM [1], jenž funguje pod záštitou společnosti Google Inc. Jde o otevřený (open source) formát, jehož myšlenkou je vybudování otevřené komunity vývojářů. Cílem vývoje kodeku bylo vytvoření efektivního kódovacího standardu pro tzv. streaming videa přes Internet. Kodek je již úspěšně implementován do služby YouTube. Kodek je zaměřen na video ve vysokém rozlišení Full High Definition (HD) 1080p (1920 x 1080) a zejména Ultra HD (4K) (3840 x 2160). Výrazné úspory bitové rychlosti dosahuje díky pokročilým predikčním algoritmům a paralelnímu kódování jednotlivých bloků obrazu.

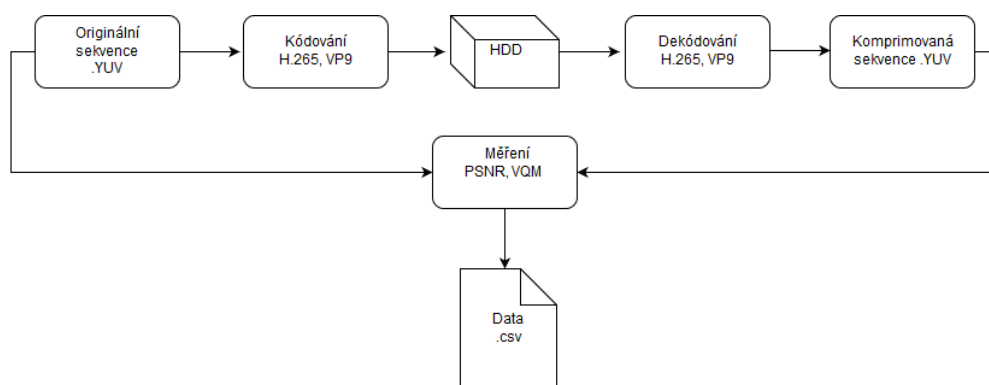
High Efficiency Video Coding (HEVC) neboli H.265 je označení pro nejnovější kódovací algoritmus, vyvíjený skupinou Joint Collaborative Team on Video Coding (JCT-VC) [2]. Jedná se o nástupce populárního kodeku s označením H.264 a disponuje oproti svému předchůdci až 50% úsporou bitového toku, zejména díky inovovaným predikčním algoritmům. Jedná se o licencovaný produkt určený pro poskytování multimediálních služeb v různých komunikačních systémech. Tento standard bude využit například pro kódování videa pozemního televizního vysílání DVB-T2 v České republice. Kromě nativní podpory rozlišení Full HD a UHD (4K) podporuje tento standard i rozlišení 8K UHD s rozlišením 8192 x 4320 pixelů.

Metody, které hodnotí kvalitu (komprimovaného) videa, se dělí do dvou základních skupin: objektivní a subjektivní.

Objektivní metriky provádí matematické vyhodnocení kvality pomocí daných algoritmů, které se liší podle konkrétní metriky. Objektivní metriky dále rozlišujeme podle toho, zda srovnávají kvalitu komprimované sekvence s původní (referenční, částečně referenční) či nikoliv (bezreferenční). Pro účely této úlohy budou použity metriky Peak Signal-to-Noise Ratio (PSNR) - referenční a Video Quality Model (VQM) – částečně referenční. V případě subjektivní metody kvalitu sekvencí hodnotí určitý počet pozorovatelů za standardizovaných podmínek. V této úloze bude použita standardizovaná metoda Absolute Category Rating (ACR).

PSNR je jednoduchá a plně referenční metrika, která vyjadřuje špičkový poměr signálu vůči šumu. Jde o velmi často využívanou metodu, mezi jejíž přednosti patří jednoduchost výpočtu. Nevýhodou této metriky je, že algoritmus výpočtu nezohledňuje vlastnosti lidského zraku. I přes její nedostatky je tato metrika často využívána jako prvotní hrubý odhad obrazové kvality. Výsledné hodnoty jsou udávány v jednotkách dB a typické výsledné hodnoty pro videa s 8-bitovou hloubkou se pohybují v rozmezí 20 – 40 dB pro nižší datové toky a 30 – 50 dB pro vysoké datové toky.

VQM je komplexní plně či částečně referenční metrika hodnocení kvality. Oproti PSNR, je tato metoda výpočetně náročnější, neboť zohledňuje model lidského vidění tzv. Human Visual System (HVS). Z toho důvodu výsledky této metody dosahují nejvyšší míry korelace ve srovnání se subjektivními testy. Tato metrika obsahuje celkem 9 různých výpočetních modelů, pro účely této úlohy bude uvažován standardizovaný model NTIA General Model. Výpočetní algoritmus tohoto modelu je založen na lineární kombinaci celkem sedmi nezávislých parametrů, které byly vybrány tak, aby co nejvíce reflektovaly zpracování obrazu HVS. Z uvedených sedmi parametrů jsou čtyři založeny na detekci změny jasové složky v obraze, dva hlídají kombinační změny chrominančních složek a poslední zaznamenává měření kontrastu a pohybu pomocí jasové složky obrazu. Výsledné hodnoty se pohybují v rozmezí od 0 (nejlepší kvalita) do 1 (nejhorší kvalita).

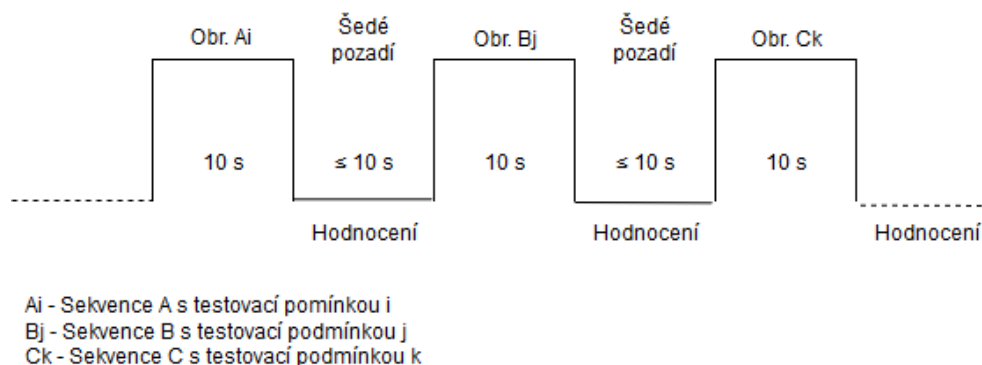


Obrázek 1 Diagram postupu hodnocení kvality sekvencí

ACR je standardizovaný (ITU BT.500-13 [3]) postup pro subjektivní hodnocení kvality videa. U této metody je pozorovatelům přehráván náhodný sled krátkých sekvencí (typicky 8 – 10 vteřinových), mezi které je vždy zařazena pauza o délce max. 10 vteřin.

Během této pauzy pozorovatel subjektivně zhodnotí kvalitu sekvence, kterou právě viděl, bez ohledu na možnost srovnání s nekomprimovanou sekvencí.

Subjektivní hodnocení pozorovatel zaznamená do dotazníku a ohodnotí známkou od 1 (nejhorší kvalita) do 5 (nejlepší kvalita). Schéma testování sekvencí metodou ACR je uvedeno na Obrázek . Z takto získaných výsledků je následně vypočítán aritmetický průměr Mean Opinion Score (MOS).



Obrázek 2 Schéma postupu testování metodou ACR

Tabulka 1 Přehled testovaných videosekvencí [4]

Název	Náhled	Popis
Tall Buildings		Statická scéna s mrakodrapy a dopravou v pozadí
Campfire Party		Dynamická noční scéna s plameny a postavami v pozadí
Fountains		Dynamická denní scéna se stříkající fontánou
Wood		Statická denní scéna v lese se slunečními paprsky

K testování jsou připraveny celkem 4 videosekvence, každá o délce 10 vteřin, v rozlišení 856 x 480 pixelů, snímkové rychlosti 30 fps, 8 – bitové barevné hloubce a formátu YUV 4:2:0. Náhled a popis jednotlivých sekvencí uvádí Tabulka 1. Tyto sekvence byly zkomprimovány oběma kodeky s nastavením bitových rychlostí: 250, 750, 1500 a 4000 kbit/s. Celkem je tedy pro testování připraveno 32 sekvencí.

## Zadání a postup měření

1. Nejprve si na připojeném TV monitoru vyzkoušíte subjektivní hodnocení kvality sekvencí metodou ACR. V kořenovém adresáři spustíte soubor „Playlist.m3u“. Na monitoru budou postupně přehrávány všechny testované sekvence v náhodném pořadí. Po zhlédnutí každé sekvence následuje pauza o délce 5 vteřin, během které bude na obrazovce šedé pozadí. Během této doby vyznačte hodnocení do přiloženého dotazníku a to na stupnici od 1 (nejhorší kvalita) do 5 (nejlepší kvalita).

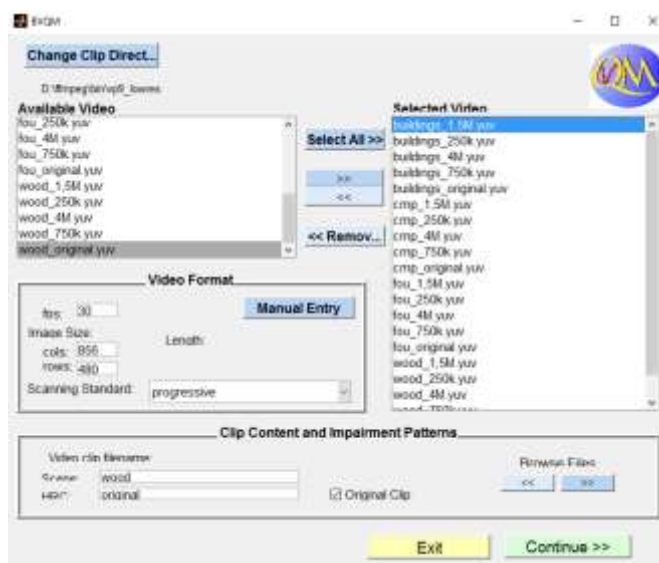
**Hodnocení provádí každý sám za sebe!** Během hodnocení přehrávání nezastavujte ani nepřetáčejte, nechte jej přehrát až do konce.

- Nyní provedete hodnocení kvality zkomprimovaných sekvencí (oběma kodeky) pomocí objektivní metriky PSNR. Za tímto účelem je ve složce Video připraven program Video Quality Measurement Tool [5][21], který se ovládá pomocí příkazové řádky s následující syntaxí:

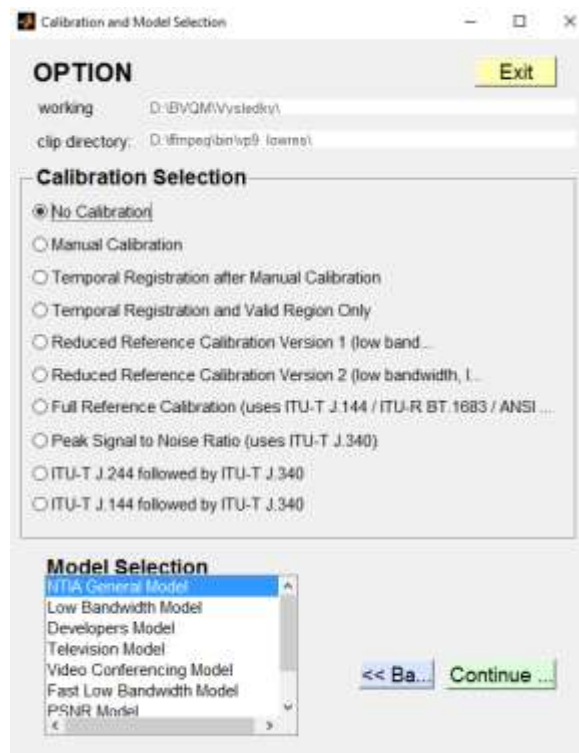
```
VQMT.exe Video\buildings_lowres_ref.yuv Video\buildings_250k.yuv 480
848 300 1 Vysledky\PSNR\buildings_250k PSNR
```

Opakované přepisování příkazu do příkazové řádky pro každou sekvenci by bylo příliš časově náročné, z toho důvodu je připraven dávkový soubor „PSNR.bat“. Po jeho spuštění začne program postupně počítat hodnotu PSNR pro každou sekvenci a výsledky zapisovat do „.csv“ souborů ve složce „Vysledky“.

- Proved'te vyhodnocení kvality sekvencí objektivní metrikou VQM. K tomuto účelu je ve složce „BVQM“ připraven program Batch Video Quality Metric [6]. Spust'te program pomocí souboru „bvqm.exe“, po načtení budete vyzváni k výběru složky, kam budou ukládány výsledky, následně vyberete složku s videosekvencemi. Nejprve je nutné vybrat sekvence pro vyhodnocení (Obrázek 3), to provedete tlačítkem „Select All >>“ V dialogovém okně je nutné ručně zadat parametry sekvencí „Enter Manually“ a vyplnit rozlišení videí a snímkovou rychlost. Nyní je možné pokračovat na výběr výpočetního modelu a kalibrace (Obrázek 4). Vzhledem k tomu, že sekvence byly pouze komprimovány, není nutné provádět kalibraci, zvolte tedy možnost „No Calibration“. V další nabídce označené „Model Selection“ vyberte požadovaný model „NTIA General Model.“ Po kliknutí na tlačítko „Continue“ začne program s výpočtem hodnot VQM. Po ukončení výpočtů je zobrazeno dialogové okno s výsledky. Je možné si prohlédnout grafickou prezentaci výsledků a poté program ukončit tlačítkem „Exit“.



Obrázek 3 Výběr souborů v programu BVQM



Obrázek 4 Nastavení parametrů pro výpočet VQM

## Zpracování výsledků

- Ad 1. Pro vyhodnocení výsledků subjektivního testu je připravena tabulka vytvořená v programu MS Excel s názvem „ACR.xlsx“. Zkontrolujte, zda je tabulka seřazena vzestupně dle sloupce „Playlist“ a pečlivě přepište do této tabulky hodnocení z vašich dotazníků. Následně celou tabulku seřadte vzestupně podle sloupce „Číslo“. Tabulka by nyní měla být seřazena podle jednotlivých sekvencí, vždy od nejnižší bitové rychlosti po nejvyšší. Z vámi uděleného skóre vypočtete aritmetický průměr a získané výsledky vyneste do sloupcových grafů, kde každý pár sloupců bude reprezentovat jednu bitovou rychlost a jeden kodek. Celkem by tedy měly vzniknout 4 grafy.
- Ad 2. Výsledky objektivní metriky PSNR jsou uloženy v jednotlivých „csv“ souborech. Tyto je nutné otevřít a najít v souboru průměrnou hodnotu každé sekvence označenou jako „average“. Všechny tyto hodnoty si poznamenejte a následně je pomocí vhodného programu (např. MS Excel) vyneste do přehledných spojnicových grafů. Osu x bude tvořit bitová rychlost jednotlivých sekvencí a osu y hodnota PSNR. Takto by měly vzniknout další 4 grafy.
- Ad 3. Výsledné hodnoty metriky VQM jsou uloženy v textovém souboru „bvqm\_summary\_report-model\_general\_YYYY\_MM\_DD@HH\_MM.txt“. Také tyto hodnoty si poznamenejte a následně zpracujte do spojnicových grafů, kde bude vynesena závislost hodnoty VQM na bitové rychlosti, pro každou sekvenci zvlášť.

## Závěr

Všechny dosažené výsledky a grafy zpracujte do přehledné zprávy o měření. Do závěru uveďte své individuální hodnocení měření. Je vhodné podrobně komentovat každý bod měření a všechny výsledné grafy. Individuální závěr by měl také uvést důležité poznatky zjištěné během měření a zodpovědět následující otázky:

- Do jaké míry korespondují výsledky vašeho subjektivního hodnocení s výsledky objektivních metrik PSNR a VQM?
- Která z objektivních metrik lépe koreluje s výsledky subjektivního hodnocení?
- Jak se liší výsledky nižších bitových toků u statické a dynamické scény?
- Stanovte pro každou sekvenci optimální bitovou rychlost, při které dosahuje vysoké obrazové kvality při co nejnižším bitovém toku.

## Kontrolní otázky

1. Jaký je rozdíl mezi objektivní a subjektivní metrikou?
2. V jakých jednotkách se uvádí výsledky PSNR?
3. Co znamená zkratka MOS?



Obrázek 5 Testovací pracoviště pro subjektivní hodnocení kvality

## Použité přístroje

- Stolní PC
- TV Samsung T24E390EW

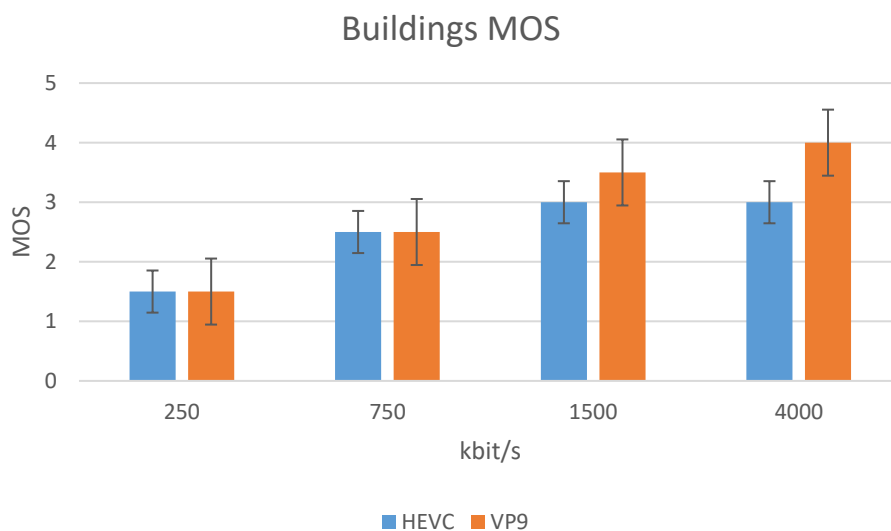
## Použitá a doporučená literatura

- [1] MUKHERJEE, D., BANKOSKI, J., GRANGE, A. *The latest open-source video codec VP9 – An overview and preliminary results*. USA: IEEE, 2013. ISBN 978-1-4799-0292-7
- [2] SULLIVAN, G. J., OHM, T., WOO-JIN, H., WIEGAND, T. *Overview of the High Efficiency Video Coding (HEVC) Standard*. *Circuits and Systems for Video Technology, IEEE Transactions on [online]*. USA: IEEE, 0121n. 1., 22(12): 1649-1668 [cit. 16.12.2015]. ISSN 1051-8215.
- [3] ITU-T Rec. BT.500 - 13. *Methodology for the subjective assessment of the quality of television pictures*. Geneva: ITU, 2012.
- [4] SONG, L., TANG, X., ZHANG, W., YANG, X., XIA, P. *The SJTU 4K Video Sequence Dataset*, the Fifth International Workshop on Quality of Multimedia Experience (QoMEX2013), Klagenfurt, Austria, 2013.
- [5] VQMT: Video Quality Measurement Tool | MMSPG. *MMSPG / EPFL* [online]. Copyright © [cit. 23.05.2017]. Dostupné z: <http://mmspg.epfl.ch/vqmt>
- [6] Description of VQM Tools - ITS. *Institute for Telecommunication Sciences - ITS* [online]. Dostupné z: <https://www.its.bldrdoc.gov/resources/video-quality-research/guides-and-tutorials/description-of-vqm-tools.aspx>
- [7] FENDRICH, V. *Srovnání standardů VP9 a H.265 pro kódování videosekvencí*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2017. 63 s. Vedoucí bakalářské práce Ing. Ladislav Polák, Ph.D.

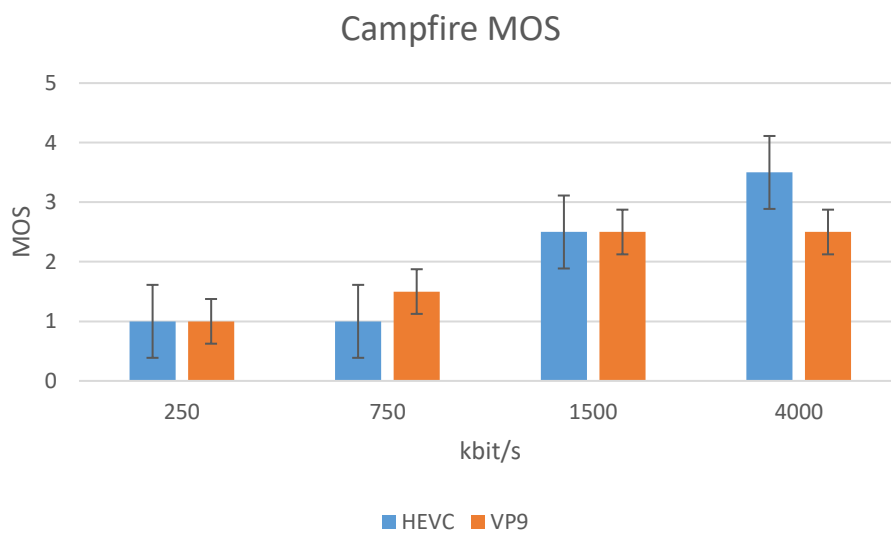


## A.2 Vzorové vypracování laboratorní úlohy

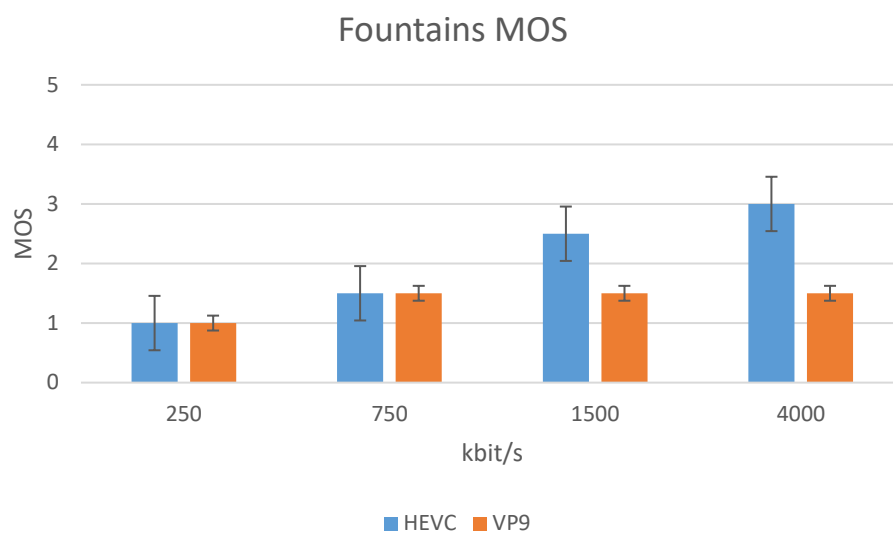
### 1. Výsledky subjektivního hodnocení metodou ACR



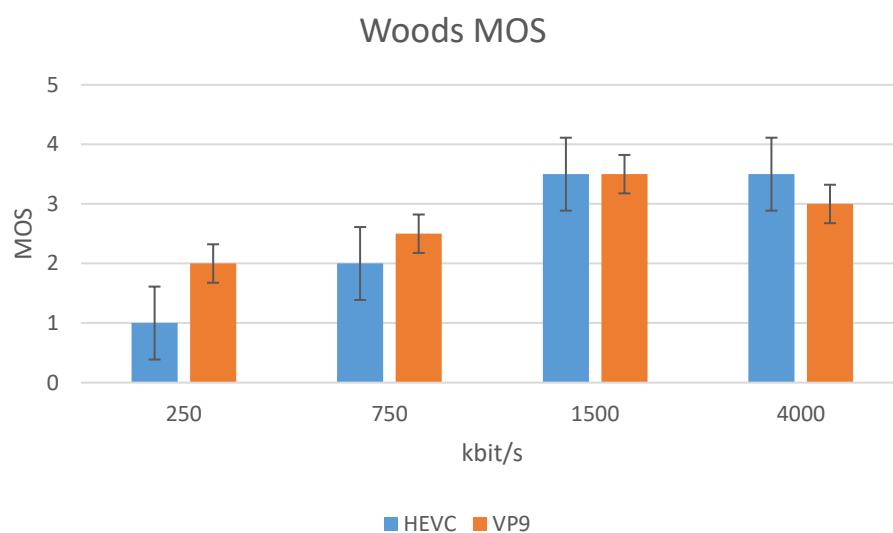
Graf č. 1 Hodnoty MOS sekvence Buildings pro jednotlivé bitové rychlosti



Graf č. 2 Hodnoty MOS sekvence Campfire pro jednotlivé bitové rychlosti

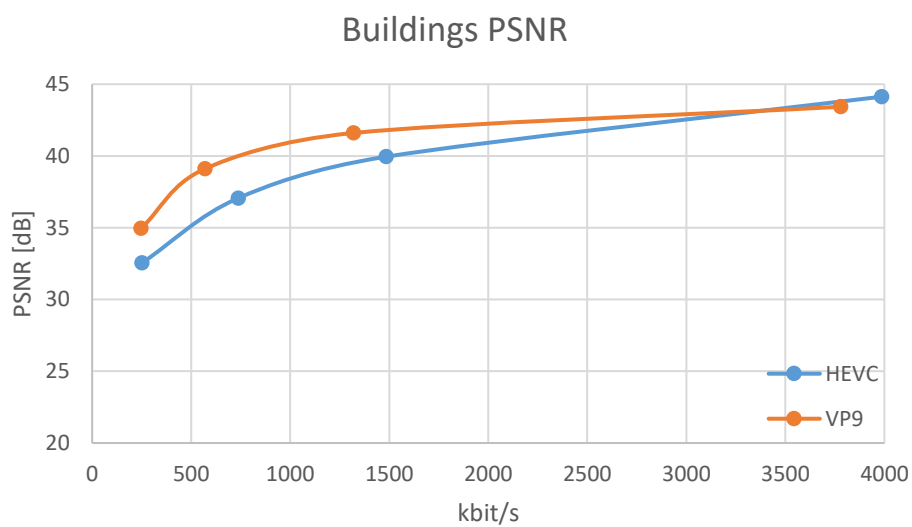


Graf č. 3 Hodnoty MOS sekvence Fountains pro jednotlivé bitové rychlosti

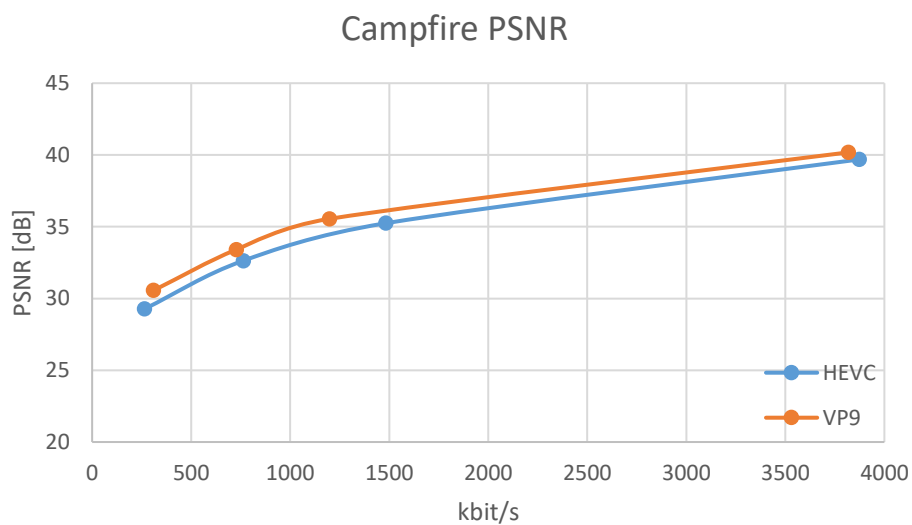


Graf č. 4 Hodnoty MOS sekvence Woods pro jednotlivé bitové rychlosti

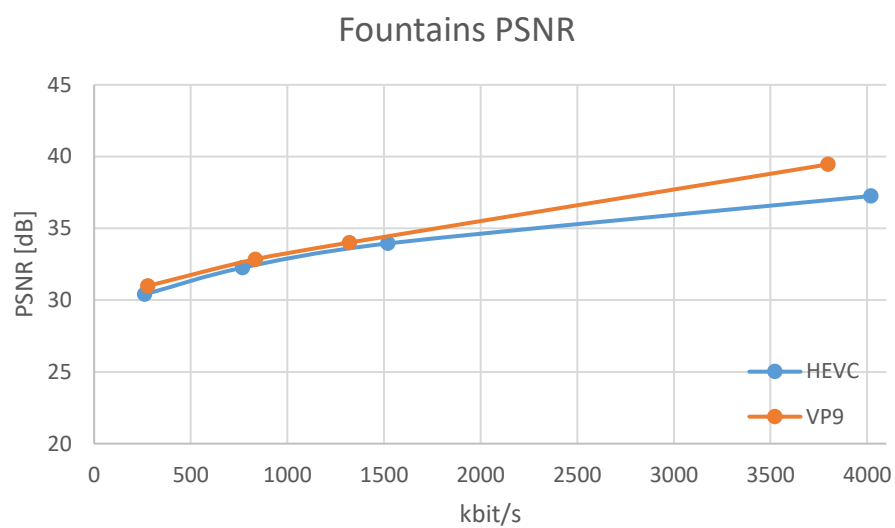
## 2. Výsledky objektivního hodnocení metrikou PSNR



Graf č. 5 Hodnoty PSNR sekvence Buildings v závislosti na bitové rychlosti



Graf č. 6 Hodnoty PSNR sekvence Campfire v závislosti na bitové rychlosti

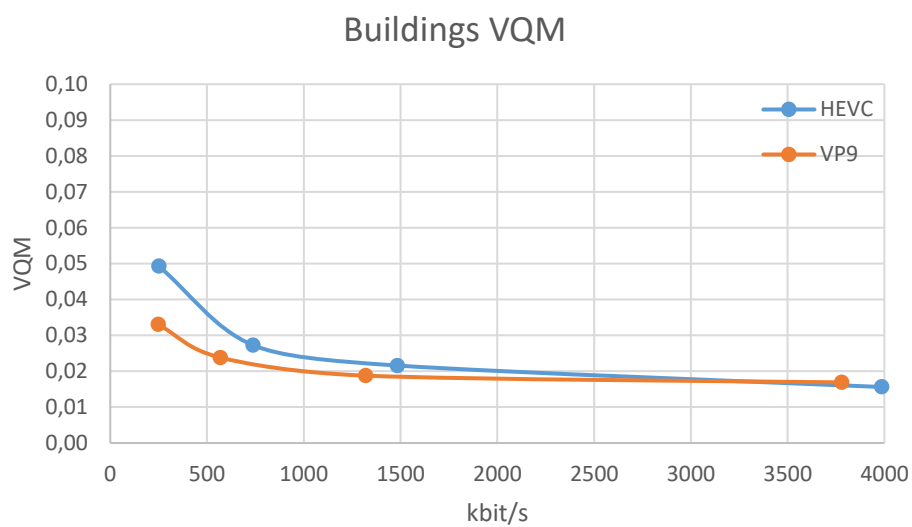


Graf č. 7 Hodnoty PSNR sekvence Fountains v závislosti na bitové rychlosti

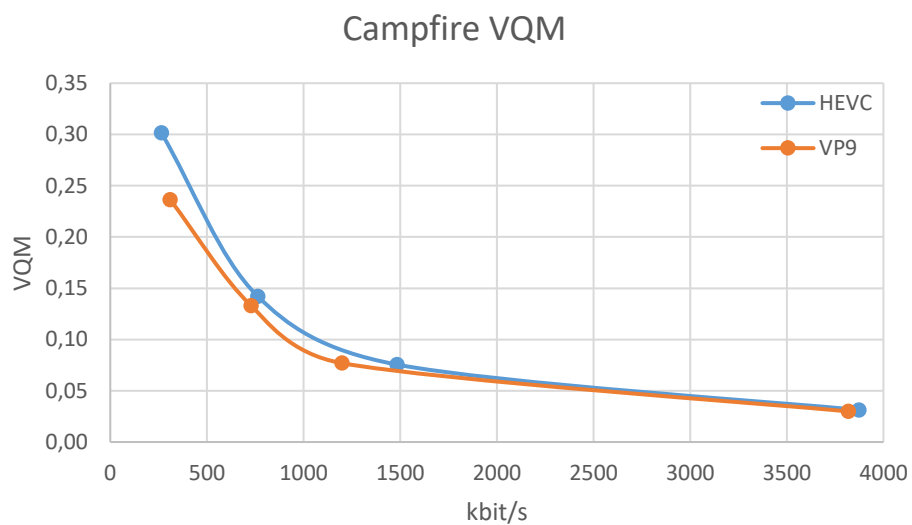


Graf č. 8 Hodnoty PSNR sekvence Woods v závislosti na bitové rychlosti

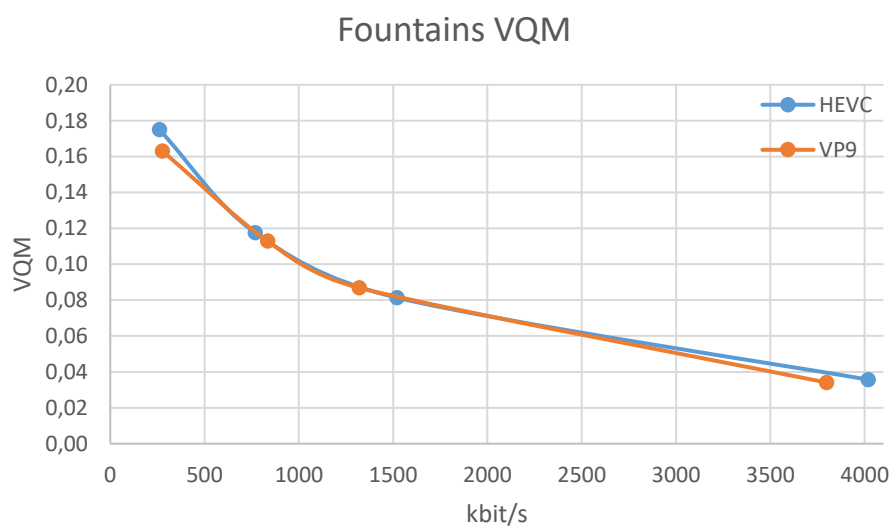
### 3. Výsledky objektivního hodnocení metrikou VQM



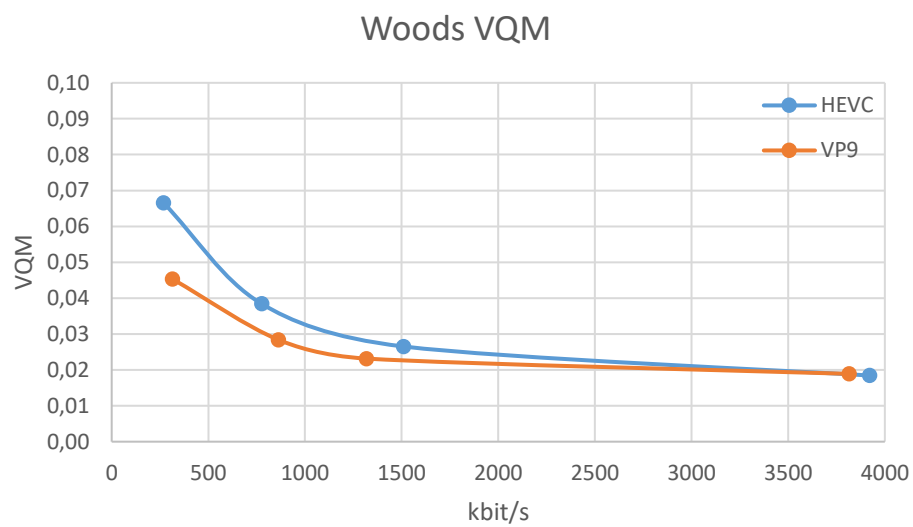
Graf č. 9 Hodnoty VQM sekvence Buildings v závislosti na bitové rychlosti



Graf č. 10 Hodnoty VQM sekvence Campfire v závislosti na bitové rychlosti



Graf č. 11 Hodnoty VQM sekvence Fountains v závislosti na bitové rychlosti



Graf č. 12 Hodnoty VQM sekvence Woods v závislosti na bitové rychlosti

## Závěr

V této laboratorní úloze jsme měli možnost se seznámit s metodami hodnocení kvality videa za účelem srovnání dvou nejnovějších kódovacích standardů VP9 a HEVC. V první části jsme provedli subjektivní hodnocení kvality metodou ACR, kdy každá sekvence byla ohodnocena na stupnici od 1 do 5. Z tohoto hodnocení byl následně vypočet aritmetický průměr MOS a vyneseno do grafů č. 1 - 4.

V další části jsme pomocí připraveného dávkového souboru provedli hodnocení objektivní metrikou PSNR a výsledné hodnoty zobrazili v grafech č. 5 – 8. Obdobným způsobem byly zpracovány také výsledky metriky VQM, která však oproti PSNR nabývá opačných hodnot (čím nižší je výsledná hodnota, tím je vyšší obrazová kvalita). Výsledky metriky VQM jsou zobrazeny v grafech č. 9 – 12.

Na výsledcích objektivních metrik PSNR i VQM lze pozorovat nárůst obrazové kvality spolu s bitovým tokem, tento trend je nejvíce patrný u dynamických scén Campfire a Fountains. U statických scén Buildings a Woods není k dosažení vysoké obrazové kvality nutný tak vysoký bitový tok, ve srovnání s dynamickou scénou.

Pro statickou scénu Buildings lze uvažovat optimální hodnotu bitového toku přibližně 1500 kbit/s, další nárůst kvality již není úměrný nárůstu bitového toku. V případě dynamické scény Campfire je k dosažení adekvátních výsledků nutný nejvyšší uvažovaný bitový tok 4000 kbit/s. Prakticky totéž platí i pro sekvenci Fountains. Pro poslední uvažovanou scénu Woods lze na základě výsledků považovat za optimální bitovou rychlost 1500 kbit/s.

## Kontrolní otázky

1. Jaký je rozdíl mezi objektivní a subjektivní metrikou?  
*Objektivní metrika označuje standardizovaný matematický algoritmus hodnocení obrazové kvality. Subjektivní metriky popisují standardizované podmínky hodnocení obrazové kvality diváky.*
2. V jakých jednotkách se uvádí výsledky PSNR?  
*dB*
3. Co znamená zkratka MOS?  
*Mean Opinion Score.*

## Použité přístroje

- Stolní PC
- TV Samsung T24E390EW

## **B DVD**

- Laboratorní úloha (LAB\...)
- Výsledky subjektivních a objektivních testů (Vysledky\...)